

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 10:14:07 ; Search time 1391.78 Seconds

(without alignments)
11437.798 Million cell updates/sec

Title: US-10-660-763-1

Perfect score: 2674
Sequence: 1 tccgggggtccgcacccgggccc.....aaaaaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2674	100.0	2674	13	US-10-003-295-1 Sequence 1, Appl1
2	2674	100.0	2674	17	US-10-660-763-1 Sequence 1, Appl1
3	2326.4	87.0	2889	13	US-10-240-965-256 Sequence 256, App
4	362.4	13.6	15297	13	US-10-003-295-3 Sequence 3, Appl1
5	362.4	13.6	15297	17	US-10-660-763-3 Sequence 3, Appl1
6	326.8	12.2	449	10	US-09-918-995-1503 Sequence 1503, Ap
7	326.8	11.1	361	9	US-09-948-802-5 Sequence 5, Appl1
8	296.2	11.1	361	15	US-10-121-925-5 Sequence 19, Appl1
9	271.6	10.2	1779	17	US-10-280-576-19 Sequence 525, App
10	201.2	7.5	3875	15	US-10-101-510-525 Sequence 1, Appl1
11	198	7.4	2955	18	US-10-384-339C-1 Sequence 1, Appl1

12	198	7.4	3370	9	US-09-967-768A-144 Sequence 144, App
13	198	7.4	3370	16	US-10-354-358-101 Sequence 101, App
14	198	7.4	3370	18	US-10-210-120-19 Sequence 19, Appl1
15	198	7.4	3370	18	US-10-776-827-82 Sequence 82, Appl1
16	198	7.4	3370	18	US-10-473-974-219 Sequence 219, App
17	185.4	6.9	3042	18	US-10-384-339C-2 Sequence 2, Appl1
18	185.4	6.9	3921	10	US-09-921-406C-29 Sequence 29, Appl1
19	185.4	6.9	3921	15	US-10-007-926A-222 Sequence 222, App
20	185.4	6.9	3921	16	US-10-269-909-19 Sequence 19, Appl1
21	185.4	6.9	3921	17	US-10-366-268-1 Sequence 1, Appl1
22	185.4	6.9	3921	17	US-10-172-118-1036 Sequence 1036, Ap
23	185.4	6.9	3921	17	US-10-342-887-1036 Sequence 1036, Ap
24	185.4	6.9	3921	17	US-10-287-226-319 Sequence 319, App
25	185.4	6.9	3921	18	US-10-648-593-1 Sequence 1, Appl1
26	185.4	6.9	3921	18	US-10-473-974-223 Sequence 223, App
27	185.4	6.9	3921	19	US-10-616-403-5 Sequence 5, Appl1
28	185.4	6.9	3935	19	US-10-897-711-1 Sequence 1, Appl1
29	176.2	6.6	1518	17	US-10-280-576-24 Sequence 24, Appl1
30	174.6	6.5	816	18	US-10-377-268-3 Sequence 3, Appl1
31	174.6	6.5	1050	18	US-10-377-268-4 Sequence 4, Appl1
32	174.6	6.5	3030	16	US-10-325-430-8 Sequence 8, Appl1
33	174.6	6.5	3416	8	US-08-987-689A-1 Sequence 1, Appl1
34	174.6	6.5	3416	15	US-10-292-524-1 Sequence 1, Appl1
35	174.6	6.5	3416	17	US-10-464-805-2 Sequence 2, Appl1
36	174.6	6.5	3416	17	US-10-305-720-1483 Sequence 1483, Ap
37	174.6	6.5	3544	14	US-10-161-803-51 Sequence 51, Appl1
38	174.6	6.5	4089	17	US-10-291-808-13 Sequence 13, Appl1
39	174.6	6.5	4089	18	US-10-620-052A-7 Sequence 7, Appl1
40	174.6	6.5	4151	14	US-10-161-803-52 Sequence 52, Appl1
41	174.6	6.5	4151	16	US-10-325-430-7 Sequence 7, Appl1
42	170.4	6.4	1611	18	US-10-476-962-3 Sequence 3, Appl1
43	170.4	6.4	1611	18	US-10-887-588-1 Sequence 1, Appl1
44	170.4	6.4	1611	18	US-10-887-588-3 Sequence 3, Appl1
45	170.4	6.4	1611	19	US-10-497-641-1 Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-003-295-1
; Sequence 1, Application US/10003295
; Publication No. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weidhu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THREEOF
; FILE REFERENCE: CL001183DIY
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-1

Query Match	100.0%	Score 2674;	DB 13;	Length 2674;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches 2674;	Conservative 0;	Indels 0;	Gaps 0;	
QY	1	TCGGGGTCCGACCGGGCTGAGTGTCCGAGCGGTCCGAGAGAGCTGCCGTGC	60	
DB	1	TCGGGGTCCGACCGGGCTGAGTGTCCGAGCGGTCCGAGAGAGCTGCCGTGC	60	
QY	61	GGAACGACACTATGGCTTCTTCTGAGCTGTGACGCCGCCAGGGGCTCTTGC	120	
DB	61	GGAACGACACTATGGCTTCTTCTGAGCTGTGACGCCGCCAGGGGCTCTTGC	120	
QY	121	AGCAATGAGAGGCGAGGCTTCTGCTGAGGAGGAGTGAAGATGATGCCAGC	180	
DB	121	AGCAATGAGAGGCGAGGCTTCTGCTGAGGAGGAGTGAAGATGATGCCAGC	180	

Db 121 AGCAAAATGAGAGGAGCCGAGCTTCCTCTA CTGAGGGGCAATGAGAAATGAGATGCCAAC 180
QY 181 GGGTCAAGATGACAGGAGATATGACAGGACTGCTTACCAATATCCCTGACAGGAGATG 240
Db 181 GGGTCAAGATGACAGGAGATATGACAGGACTGCTTACCAATATCCCTGACAGGAGATG 240
QY 241 GGGGCGAGAGCCGGGCGCATCAGCCCTTGACAGCCCATCATGATCAGTCCCTGGCTGATGATCA 300
Db 241 GGGGCGAGAGCCGGGCGCATCAGCCCTTGACAGCCCATCATGATCAGTCCCTGGGCTGATGATCA 300
QY 301 CCAGCCAACTGAGGGGCTGAGGCGCTTGCTGCGGACGACCGAGAGATCTGAACTCAG 360
Db 301 CCAGCCAACTGAGGGGCTGAGGCGCTTGCTGCGGACGACCGAGAGATCTGAACTCAG 360
QY 361 GGGCCCTGAGCAAGCTGAGGCTGCTCATCCGGGAGAGGGGAGGAGCTTCGAAAGCTTACA 420
Db 361 GGGCCCTGAGCAAGCTGAGGCTGCTCATCCGGGAGAGGGGAGGAGCTTCGAAAGCTTACA 420
QY 421 GCGAGAGTGGACGACGCTGACAGAGAGCTCACAAAGCCACAGCCAGACATTTGAGA 480
Db 421 GCGAGAGTGGACGACGCTGACAGAGAGCTCACAAAGCCACAGCCAGACATTTGAGA 480
QY 481 AGCTGAAGAGCCAGTACCGAGCTCTGCGCACGGGACAGTGCCTCAAGCCAGCCAGATCC 540
Db 481 AGCTGAAGAGCCAGTACCGAGCTCTGCGCACGGGACAGTGCCTCAAGCCAGCCAGATCC 540
QY 541 AGGAGGCCAGCAAGACCAAGGACCGTGAACAGGSCCAAGGACAGTATGTCGCGAGCTGT 600
Db 541 AGGAGGCCAGCAAGACCAAGGACCGTGAACAGGSCCAAGGACAGTATGTCGCGAGCTGT 600
QY 601 GGAAGCTCTTTGCTCACCAACCGCTATGTCGCGGCTGCGGGCTGCGGAGCTACACC 660
Db 601 GGAAGCTCTTTGCTCACCAACCGCTATGTCGCGGCTGCGGGCTGCGGAGCTACACC 660
QY 661 ACCAGACCAACCAACGAGCTCTGCTGCGCGGCTGCTGCGGCTGCTGCGGAGCTTCAACG 720
Db 661 ACCAGACCAACCAACGAGCTCTGCTGCGCGGCTGCTGCGGCTGCTGCGGAGCTTCAACG 720
QY 721 AGGAGATGGCTTGCATCTGAAAGAGATCGTGAAGGATATCTGAGGATTTAGAGGCTGG 780
Db 721 AGGAGATGGCTTGCATCTGAAAGAGATCGTGAAGGATATCTGAGGATTTAGAGGCTGG 780
QY 781 TGCAGATGAGTGTGTGAGGCTTACCGGAGATGAGCTGAGCTGTCGCCGATCAG 840
Db 781 TGCAGATGAGTGTGTGAGGCTTACCGGAGATGAGCTGAGCTGTCGCCGATCAG 840
QY 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGTCCGCACTGACCTGACCTCC 900
Db 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGTCCGCACTGACCTGACCTCC 900
QY 901 GTGTCACTTTCGATGAGTCACTGCTTGAAGGAGGTGAACCGCTGAGGCTGGGAGCTCC 960
Db 901 GTGTCACTTTCGATGAGTCACTGCTTGAAGGAGGTGAACCGCTGAGGCTGGGAGCTCC 960
QY 961 AGCTGAACGAGCTGATGTGAGAGAGGTGACAGACGCTGACCTGATGACAGATGAG 1020
Db 961 AGCTGAACGAGCTGATGTGAGAGAGGTGACAGACGCTGACCTGATGACAGATGAG 1020
QY 1021 TGGCTGTGGCCACCGAGATGTGTTCAGGCGGACAGAGATGTTATGCGAGCTGCAACAG 1080
Db 1021 TGGCTGTGGCCACCGAGATGTGTTCAGGCGGACAGAGATGTTATGCGAGCTGCAACAG 1080
QY 1081 AGCTCGGAAATGAAGAGGAGAACCAACCCCGGGAGCGGGGTGAGCTGTGGGAAAG 1140
Db 1081 AGCTCGGAAATGAAGAGGAGAACCAACCCCGGGAGCGGGGTGAGCTGTGGGAAAG 1140
QY 1141 GCGAAGTGTGCAAGAGCACTGACAGGAGCTGAGAGTATGCGCTGACAGCCAGC 1200
Db 1141 GCGAAGTGTGCAAGAGCACTGACAGGAGCTGAGAGTATGCGCTGACAGCCAGC 1200
QY 1201 TGCAGGCCACGACGAGAGTGTCTGACAGCAAGCTGAGACCTTGGGCGCGGAGCCCC 1260
Db 1201 TGCAGGCCACGACGAGAGTGTCTGACAGCAAGCTGAGACCTTGGGCGCGGAGCCCC 1260

QY 1261 CGCCTGTGTCTCTCTGACAGATGACCGGCACTCCACGTGCTCTGTGAGACAGAGCAG 1320
Db 1261 CGCCTGTGTCTCTCTGACAGATGACCGGCACTCCACGTGCTCTGTGAGACAGAGCAG 1320
QY 1321 AGGGGAGAGAGACCCACGCTGAGATCTTTAAGACCAATCTCAGAAATCTTCGCC 1380
Db 1321 AGGGGAGAGAGACCCACGCTGAGATCTTTAAGACCAATCTCAGAAATCTTCGCC 1380
QY 1381 CCAAGTCTCGAACCTGTACCGATCGGAGAGGGGAGGCTTCTCTAGCATTCCTTGTCTCA 1440
Db 1381 CCAAGTCTCGAACCTGTACCGATCGGAGAGGGGAGGCTTCTCTAGCATTCCTTGTCTCA 1440
QY 1441 TGCACCACTTATGAGACCCACAGCCCTTACCAAGAGAGTGTGTCTGACAGA 1500
Db 1441 TGCACCACTTATGAGACCCACAGCCCTTACCAAGAGAGTGTGTCTGACAGA 1500
QY 1501 GGGCTGTGCCAAGGACAAAGTGGTGTGAAACCATAGAGACCTGTGTGGTGTGACAGA 1560
Db 1501 GGGCTGTGCCAAGGACAAAGTGGTGTGAAACCATAGAGACCTGTGTGGTGTGACAGA 1560
QY 1561 TTGAGACGGGAGAACTTTGGCGAAGTTCAGGGGAGCGCTGGGAGCGGACCAACCCCTGG 1620
Db 1561 TTGAGACGGGAGAACTTTGGCGAAGTTCAGGGGAGCGCTGGGAGCGGACCAACCCCTGG 1620
QY 1621 TGGCGGTGAAGTCTTGTGAGAGAGCGCTCCACCTGACCTCAAGGCGCAAGTTTCTACAG 1680
Db 1621 TGGCGGTGAAGTCTTGTGAGAGAGCGCTCCACCTGACCTCAAGGCGCAAGTTTCTACAG 1680
QY 1681 AAGGAGATCTTGAAGCAGTACAGCCACCCCAACATGTCGCTCATTTGTGTCTGCA 1740
Db 1681 AAGGAGATCTTGAAGCAGTACAGCCACCCCAACATGTCGCTCATTTGTGTCTGCA 1740
QY 1741 CCCAAGAGAGCCCAATCTACATGTCATGAGACCTTGTGAGGGGGGCGACTTCCTGACCT 1800
Db 1741 CCCAAGAGAGCCCAATCTACATGTCATGAGACCTTGTGAGGGGGGCGACTTCCTGACCT 1800
QY 1801 TCTCTCGACAGAGGGGGGCCCTGCGGGTGAAGACTGTGTCAGATGATGAGGAGATG 1860
Db 1801 TCTCTCGACAGAGGGGGGCCCTGCGGGTGAAGACTGTGTCAGATGATGAGGAGATG 1860
QY 1861 CAGCTGTGGCATGAGATCTGAGAGCAGATGTGCTGACACCGGAGACTGGCTGCTC 1920
Db 1861 CAGCTGTGGCATGAGATCTGAGAGCAGATGTGCTGACACCGGAGACTGGCTGCTC 1920
QY 1921 GGAATCGCTGGTGAACAAGAGAAATGCTGTAAGTCAATGTAATTTGGAGTGTCCGAG 1980
Db 1921 GGAATCGCTGGTGAACAAGAGAAATGCTGTAAGTCAATGTAATTTGGAGTGTCCGAG 1980
QY 1981 AGGAAAGCCGATGGGCTCTATGAGCCTCAGGGGCGCTCAGACAAAGTCCCGTGAATGGA 2040
Db 1981 AGGAAAGCCGATGGGCTCTATGAGCCTCAGGGGCGCTCAGACAAAGTCCCGTGAATGGA 2040
QY 2041 CCGACCTGAGGCGCTTAACTAACCGGCGTACTCTCCGAAAGCGAGCTGTGAGGCTTTG 2100
Db 2041 CCGACCTGAGGCGCTTAACTAACCGGCGTACTCTCCGAAAGCGAGCTGTGAGGCTTTG 2100
QY 2101 GGAATCTGTCTGAGAGACCTTCAAGCTGGGGGCGCTCCCTATCCCAACTCAGAAATC 2160
Db 2101 GGAATCTGTCTGAGAGACCTTCAAGCTGGGGGCGCTCCCTATCCCAACTCAGAAATC 2160
QY 2161 AGCAGACAGGAGATTTGTGAGAAAGGGGGGCGCTGTGCTGCGCCAGAGCTGTCTG 2220
Db 2161 AGCAGACAGGAGATTTGTGAGAAAGGGGGGCGCTGTGCTGCGCCAGAGCTGTCTG 2220
QY 2221 ATGCCGTGTTACGCTCATGAGACAGTGTGAGGCTGTAGGCTGTGGAGCCAGCT 2280
Db 2221 ATGCCGTGTTACGCTCATGAGACAGTGTGAGGCTGTAGGCTGTGGAGCCAGCT 2280
QY 2281 TGCAGCACTTACCAAGAGCTGACAGAGATCCGAAAGGGGCAATCGGTGAGGCTGGGACC 2340
Db 2281 TGCAGCACTTACCAAGAGCTGACAGAGATCCGAAAGGGGCAATCGGTGAGGCTGGGACC 2340

[illegible]

```

RESULT 2
US-10-660-763-1
; Sequence 1, Application US/10660763
; Publication No. US20040063130A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weidun et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01183DIV1
; CURRENT APPLICATION NUMBER: US/10/660,763
; CURRENT FILING DATE: 2003-09-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-660-763-1

```

Query Match	100.0%;	Score 2674;	DB 17;	Length 2674;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2674;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	TCGGGGGTCC	CA	CGGGCC	CT	GA	GTGG	AT	CCG	AG	GC	CG	CG	GC	60
Db	1	TCGGGGGTCC	CA	CGGGCC	CT	GA	GTGG	AT	CCG	AG	GC	CG	CG	GC	60
QY	61	GGAA	CAG	CA	CT	AT	G	GG	CT	T	T	CT	T	CT	GA
Db	61	GGAA	CAG	CA	CT	AT	G	GG	CT	T	T	CT	T	CT	GA
QY	121	AGCA	AAT	GC	AG	AG	GC	CG	AG	CT	T	CG	T	CT	AC
Db	121	AGCA	AAT	GC	AG	AG	GC	CG	AG	CT	T	CG	T	CT	AC
QY	181	GGGT	CA	AG	AG	GT	AT	GA	CG	AG	CA	GT	AT	GA	CG
Db	181	GGGT	CA	AG	AG	GT	AT	GA	CG	AG	CA	GT	AT	GA	CG
QY	241	GGGG	CC	AG	GC	CG	GG	CC	CA	T	CA	GC	CT	GA	CG
Db	241	GGGG	CC	AG	GC	CG	GG	CC	CA	T	CA	GC	CT	GA	CG
QY	301	CCAG	CC	AA	A	CT	GA	GG	GC	CT	GA	GC	CT	GA	GC
Db	301	CCAG	CC	AA	A	CT	GA	GG	GC	CT	GA	GC	CT	GA	GC

QY	36	GGCCCCGAGCAACCTGAGCTGCAATCCGGGAAACGGACAGACTTTCGAAGACTTACA	420
Dp	361	GGCCCCGAGCAACCTGAGCTGCAATCCGGGAAACGGACAGACTTTCGAAGACTTACA	420
QY	421	GCGAGCAGTGGCAGACCTGACAGGAGAGCTCAACCAAGACCAGACGACAGCATTTGAGA	480
Dp	421	GCGAGCAGTGGCAGACCTGAGAGGAGAGCTCAACCAAGACCAGACGACAGCATTTGAGA	480
QY	481	AGCTGAAGACCACTACCTGAGCTCTGCGACGGGACAGTGTCCCAAGCCAGCAATAC	540
Dp	481	AGCTGAAGACCACTACCTGAGCTCTGCGACGGGACAGTGTCCCAAGCCAGCAATAC	540
QY	541	AGGAGGCGACGAAGAAGACGACCTGTACAAAGGCCAAGACCAACTATGTGGCGACCTGT	600
Dp	541	AGGAGGCGACGAAGAAGACGACCTGTACAAAGGCCAAGACCAACTATGTGGCGACCTGT	600
QY	601	GGAAGCTCTTGTGTCAACCAACCGCTATGTGTGGGCGTGGGGCTTGGCAGACTTAC	660
Dp	601	GGAAGCTCTTGTGTCAACCAACCGCTATGTGTGGGCGTGGGGCTTGGCAGACTTAC	660
QY	661	ACGACGACCAACCAACGACTCTCTGTGCTCGGCGCTGTGCGGTCACCTGACAGACTTGCAC	720
Dp	661	ACGACGACCAACCAACGACTCTCTGTGCTCGGCGCTGTGCGGTCACCTGACAGACTTGCAC	720
QY	721	AGGAGATGCGTGTGATCTCTGAAGAGATCTCTGCGAGAAATACCTGGAATTTAGCACTTGG	780
Dp	721	AGGAGATGCGTGTGATCTCTGAAGAGATCTCTGCGAGAAATACCTGGAATTTAGCACTTGG	780
QY	781	TGCAAGATGAGAGTGTGGCCATTCAACGGGGAGATGGCTGCAGCTGTGGCCGCAATCCAGC	840
Dp	781	TGCAAGATGAGAGTGTGGCCATTCAACGGGGAGATGGCTGCAGCTGTGGCCGCAATCCAGC	840
QY	841	CTGAGGCTGAGTACCAAGGCTTCTCTGCGACAGTATGGGTCGCACTTGCACCTCCACCT	900
Dp	841	CTGAGGCTGAGTACCAAGGCTTCTCTGCGACAGTATGGGTCGCACTTGCACCTCCACCT	900
QY	901	GTCGTACAGCTTCAGATGATCACTGTTTGAAGAGGGTGAACCGCTGAGCTTGGGGAGCTTC	960
Dp	901	GTCGTACAGCTTCAGATGATCACTGTTTGAAGAGGGTGAACCGCTGAGACCTTGGGGAGCTTC	960
QY	961	AGCTGAACGAGCTACCTGTGGAAGGGGTGCGACACAGCTGACTCAGTGAAGATGAGC	1020
Dp	961	AGCTGAACGAGCTACCTGTGGAAGGGGTGCGACACAGCTGACTCAGTGAAGATGAGC	1020
QY	1021	TGGCTGTGGCCACCGAGATGGTGTTCAGGGCGGCAGAGATGGTTACCGAGCTGCAACAG	1080
Dp	1021	TGGCTGTGGCCACCGAGATGGTGTTCAGGGCGGCAGAGATGGTTACCGAGCTGCAACAG	1080
QY	1081	AGCTCCGGAAATGAAGAGAGAACCCACCCCGGGAGCGGGTGCACCTGCTGGGCAAGA	1140
Dp	1081	AGCTCCGGAAATGAAGAGAGAACCCACCCCGGGAGCGGGTGCACCTGCTGGGCAAGA	1140
QY	1141	GGCAAGTGTGCAAGACCTGACGAGGGGCTGACAGTATGGCTGTGACGCCAGGCGCAAGC	1200
Dp	1141	GGCAAGTGTGCAAGACCTGACGAGGGGCTGACAGTATGGCTGTGACGCCAGGCGCAAGC	1200
QY	1201	TGCAAGGCCCAAGAGAAATTTCTGCAAGCAAGCTGAGACACTGTGGCCCGGCGAGCCCC	1260
Dp	1201	TGCAAGGCCCAAGAGAAATTTCTGCAAGCAAGCTGAGACACTGTGGCCCGGCGAGCCCC	1260
QY	1261	CGCTGTGTGCTCTCTGACAGATGACCGCACTCCACGTGCTCTGAGAGCAGAGCGAG	1320
Dp	1261	CGCTGTGTGCTCTCTGACAGATGACCGCACTCCACGTGCTCTGAGAGCAGAGCGAG	1320
QY	1321	AGGGGGGAAGAAGACCCAGCTGTGAAGATCTTTAAGACCAATTCAGAGAAATCTTCCGCC	1380
Dp	1321	AGGGGGGAAGAAGACCCAGCTGTGAAGATCTTTAAGACCAATTCAGAGAAATCTTCCGCC	1380
QY	1381	CCAAGTTCAGAACTGTACCGACTGGAAGGGAGAGGCTTTCCTAGAGATTCCTTTGCTCA	1440
Dp	1381	CCAAGTTCAGAACTGTACCGACTGGAAGGGAGAGGCTTTCCTAGAGATTCCTTTGCTCA	1440
QY	1441	TCGACCAACTTACTGAGACCCCAAGCAGCCCTTCAACCAAGAAAGTGTGTCTGTGACA	1500

421 GCGAGCAGTGGAGCAGCTGACAGGAGCTCAACAAAGCCACAGCCAGAGCATTTGAGA 480
|||||
538 GCGAGCAGTGGAGCAGCTGACAGGAGCTCAACAAAGCCACAGCCAGAGCATTTGAGA 597
|||||
481 AGCTGAGAGCCAGATACCAGAGCTTGGCA GGGGACAGTGCACCAAGCCCAAGTACC 540
598 AGCTGAGAGCCAGATACCAGAGCTTGGCA GGGGACAGTGCACCAAGCCCAAGTACC 657
541 AGGAGGCCAGAAAGCAAGAGACCTGTACAAAGGCCAAGACAAATATGTGCGAGCTGT 600
658 AGGAGGCCAGAAAGCAAGAGACCTGTACAAAGGCCAAGACAAATATGTGCGAGCTGT 717
601 GGAAGCTCTTGTCTCAACCAACCGCTATGTGTGGGCGTGGGGGTGGAGCTTACACC 660
718 GGAAGCTCTTGTCTCAACCAACCGCTATGTGTGGGCGTGGGGGTGGAGCTTACACC 777
661 ACCAGACACCAACCAACAGCTCTGTGCGGCGGCGTGTGAGTCACTGACAGACCTGACAG 720
778 ACCAGACACCAACCAACAGCTCTGTGCGGCGGCGTGTGAGTCACTGACAGACCTGACAG 837
721 AGGAGATGGCTTGCATCTGTAAAGAGATCTCTGACAGAAATACCTGAGATTTAGCAGCTTG 780
838 AGGAGATGGCTTGCATCTGTAAAGAGATCTCTGACAGAAATACCTGAGATTTAGCAGCTTG 897
781 TGCAGGATGAGGATGGTGGCATTTCAACGGGAGATGGCTGAGTGTGCGGCGATCCAGC 840
898 TGCAGGATGAGGATGGTGGCATTTCAACGGGAGATGGCTGAGTGTGCGGCGATCCAGC 957
841 CTGAGGCTAGTATCCAAAGGCTTCTGTGACAGATATGGTTCGCAACCTGACGTGCCACCT 900
958 CTGAGGCTAGTATCCAAAGGCTTCTGTGACAGATATGGTTCGCAACCTGACGTGCCACCT 1017
901 GTGTCACTGTTGCAATGATCATGTCTTGAAGAGGGTGAACCGCTGAGCTGTGGGAGCTTC 960
1018 GTGTCACTGTTGCAATGATCATGTCTTGAAGAGGGTGAACCGCTGAGCTGTGGGAGCTTC 1077
961 AGCTAAACGAGCTGACTGTGAGAGAGGTGACACACGCTGACCTTACAGTACAGATGAGAC 1020
1078 AGCTAAACGAGCTGACTGTGAGAGAGGTGACACACGCTGACCTTACAGTACAGATGAGAC 1137
1021 TGGCTGTGGCCACCGAGATGTGTGTTACAGGCGGACAGAGATGGTTACGACGCTGCAACAG 1080
1138 TGGCTGTGGCCACCGAGATGTGTGTTACAGGCGGACAGAGATGGTTACGACGCTGCAACAG 1197
1081 AGCTTCGGAATGAAGAGAGAAACACCCACCTCCGGAGCGGGTGAAGCTGTGGGCAAGA 1140
1198 AGCTTCGGAATGAAGAGAGAAACACCCACCTCCGGAGCGGGTGAAGCTGTGGGCAAGA 1257
1141 GGCAGATGTGTGCAAGAGACATGTGACAGGGGTGCAAGTACGCTGTGACACCGCCAAAGC 1200
1258 GGCAGATGTGTGCAAGAGACATGTGACAGGGGTGCAAGTACGCTGTGACACCGCCAAAGC 1317
1201 TGCAGGCGCCAGCAGAGAGTGTGTCAGAACCAAGCTGAGACACTTGGGCGCCGAGAGCCCC 1260
1318 TGCAGGCGCCAGCAGAGAGTGTGTCAGAACCAAGCTGAGACACTTGGGCGCCGAGAGCCCC 1377
1261 GCGCTGTGCTGTCTCTGACAGATGACCGGCACTCCAGTGTGTCTTGGAGACAGAGCGAG 1320
1378 GCGCTGTGCTGTCTCTGACAGATGACCGGCACTCCAGTGTGTCTTGGAGACAGAGCGAG 1437
1321 AGGGGGGAAAGACACCCACAGCTGAGAGATCTTTAAGAGCCACATCTTACAGGAATCTTCCGCC 1380
1438 AGGGGGGAAAGACACCCACAGCTGAGAGATCTTTAAGAGCCACATCTTACAGGAATCTTCCGCC 1497
1381 CCAAGTTCTCG----- 1391
1498 CCAAGTTCTCGCTCCCTCCACCGCTGCACTATTCGAGAGTGGAGAGAGCCCTGTCATG 1557
1392 ----- 1391
1558 AGCAGCTGTATACACGAGGAGCAATCCGAGGGCAGAGAGTGGCTGAGCTGTGTGACT 1617
1392 ----- 1391

1618 CTGGGACATTTCTGTGTGGGGAGAGACCGAGGCAAGAGAGTACGTGTGTGTCTGT 1677
1392 -----AACCTTACCGACTGGAAG 1410
1678 GGGATGTGTGCCCCGGGCACTTATATCATCAAGTCTTGGATTAACCTGTACCACTGGAAG 1737
1411 GGGAGGCTTTTCTTACATTTCTTGTCTCATGACCACTACTATTGAGACCCGAGGCCCC 1470
1738 GGGAGGCTTTTCTTACATTTCTTGTCTCATGACCACTACTATTGAGACCCGAGGCCCC 1797
1471 TCACCAAGAAAGAGTGTGTGTCTGACAGAGGCTGTGCGCCAAAGACAAAGTGGTGTGA 1530
1798 TCACCAAGAAAGAGTGTGTGTCTGACAGAGGCTGTGCGCCAAAGACAAAGTGGTGTGA 1857
1531 ACCATGAGAGACTGTGTGTGTGGTGAAGACAGATTTGAACGGGGAACCTTTGGCAAGTGTTC 1590
1858 ACCATGAGAGACTGTGTGTGTGGTGAAGACAGATTTGAACGGGGAACCTTTGGCAAGTGTTC 1917
1591 GGGAGGCTTGTGAGACCCGACAAACCTCTGTGTGGCGGTGAAGTCTTGTGAGAGAGCTTC 1650
1918 GGGAGGCTTGTGAGACCCGACAAACCTCTGTGTGGCGGTGAAGTCTTGTGAGAGAGCTTC 1977
1651 CACTGACCTCAAGGCGCAAGTTTCTACAGGAAGCAGATCTGAAAGCAGTACAGGCAAC 1710
1978 CACTGACCTCAAGGCGCAAGTTTCTACAGGAAGCAGATCTGAAAGCAGTACAGGCAAC 2037
1711 CCAACATGTGTGTCTCATTTGTGTGTGCAACCAAGAGCCCACTTACATGTGTATG 1770
2038 CCAACATGTGTGTCTCATTTGTGTGTGCAACCAAGAGCCCACTTACATGTGTATG 2097
1771 AGCTTGTGAGGGGGGCGCACTTCTGACCTTCTCTGCGACGAGAGGGGCGCGCTGCGGG 1830
2098 AGCTTGTGAGGGGGGCGCACTTCTGACCTTCTCTGCGACGAGAGGGGCGCGCTGCGGG 2157
1831 TGAAGACTCTGTGCAAGATGTGTGGGGGATGACGCTGTGCGACATGAGTACTGAGAGGA 1890
2158 TGAAGACTCTGTGCAAGATGTGTGGGGGATGACGCTGTGCGACATGAGTACTGAGAGGA 2217
1891 AGTGTGATTCACCGGAGCCTGTGTCTGCGAATGCTGTGTGAACAGAGAAATGTCTC 1950
2218 AGTGTGATTCACCGGAGCCTGTGTCTGCGAATGCTGTGTGAACAGAGAAATGTCTC 2277
1951 TGAAGTCAATGATCTTTGGGATGTCTCCGAGAGAAAGCCGATGGGCTTATGCAAGCTCAG 2010
2278 TGAAGTCAATGATCTTTGGGATGTCTCCGAGAGAAAGCCGATGGGCTTATGCAAGCTCAG 2337
2011 GGGGCTCTGACAAAGTCCCGGAAAGTGAACGCACTGAGGCGCTTAACTTACGAGCGCT 2070
2338 GGGGCTCTGACAAAGTCCCGGAAAGTGAACGCACTGAGGCGCTTAACTTACGAGCGCT 2397
2071 ACTCTCCGAAAGCGACGATGTGAGACTTGTGAGCTTGTGAGAGACTTCAAGCTTG 2130
2398 ACTCTCCGAAAGCGACGATGTGAGACTTGTGAGCTTGTGAGAGACTTCAAGCTTG 2457
2131 GGGCTCTCCCTTATCCAACTTCAAGCAATCAGCAACACGAGAGATTTGTGAGAGAGGGG 2190
2458 GGGCTCTCCCTTATCCAACTTCAAGCAATCAGCAACACGAGAGATTTGTGAGAGAGGGG 2517
2191 GCGGTCTGCGCCGACAGAGCTGTGTCTGATGCGGTGTGAGGCTCATGAGAGAGTGTCT 2250
2518 GCGGTCTGCGCCGACAGAGCTGTGTCTGATGCGGTGTGAGGCTCATGAGAGAGTGTCT 2577
2251 GGGCTTATGAGCTGTGGGACGCGCCAGCTTTCAGACCACTTACAGAGAGCTGACAGAGA 2310
2578 GGGCTTATGAGCTGTGGGACGCGCCAGCTTTCAGACCACTTACAGAGAGCTGACAGAGA 2637
2311 TCCGAAAGCGGATGTGTGAGGCTGAGACCCCTTCTCAAGCTGTGTGAGGCTTGTGAGGCTC 2370
2638 TCCGAAAGCGGATGTGTGAGGCTGAGACCCCTTCTCAAGCTGTGTGAGGCTTGTGAGGCTC 2697
2371 TAGGTGAGCTCTCAAGGCGCTCAAGCTCAATGCTGACAGCTTTCACAGTCTGAGACT 2430

```
Db 2698 TAGGTGACGCTCTCTGAGCGGCTCCAGCTCATATGTGACAGCTTTCACAGCTCTGGAAT 2757
Qy 2431 CCTGCCACGATTCACACTGCGGCGAGATGACAGCGCGCTCTCTCTGTGTCTCTG 2490
Db 2758 CCTGCCACGATTCACACTGCGGCGAGATGACAGCGCGCTCTCTCTGTGTCTCTG 2817
Qy 2491 CTGCTCCAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2550
Db 2818 CTGCTCCAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2877
Qy 2551 AAAAAAAAAA 2560
Db 2878 AAAAAAAAAA 2887
```

```
RESULT 4
US-10-003-295-3
; Sequence 3, Application US/10003295
; Publication No. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-3
```

```
Query Match 13.6%; Score 362.4; DB 13; Length 15297;
Best Local Similarity 99.7%; Pred. No. 1.2e-46;
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 2187 GGGGGGCGTCTGCTCTGCTCCAGAGCTGTGTCTGATGCGGTGTTCAAGCTCATGAGCAG 2246
Db 12938 GGGGGGCGTCTGCTCTGCTCCAGAGCTGTGTCTGATGCGGTGTTCAAGCTCATGAGCAG 12997
Qy 2247 TGCTGGGCTATGAGGCTGGGCGAGGCGCCAGCTTACAGCAATACAGAGAGCTGCAG 2306
Db 12998 TGCTGGGCTATGAGGCTGGGCGAGGCGCCAGCTTACAGCAATACAGAGAGCTGCAG 13057
Qy 2307 AGCATCCGAAAGCGGATCGGTGAGGCTGGGACCCCTTCTCAAGCTGTGAGCTCTGCA 2366
Db 13058 AGCATCCGAAAGCGGATCGGTGAGGCTGGGACCCCTTCTCAAGCTGTGAGCTCTGCA 13117
Qy 2367 GGGCTTGGTGGAGCTCTCTCAAGCGGCTTCAGACTATATGCTGACAGCTTTACAGTCTG 2426
Db 13118 GGGCTTGGTGGAGCTCTCTCAAGCGGCTTCAGACTATATGCTGACAGCTTTACAGTCTG 13177
Qy 2427 GACTCTGACACAGATCCACATGCGGCGAGAGAGAGCGCGCTGCTCTCTGCTGTC 2486
Db 13178 GACTCTGACACAGATCCACATGCGGCGAGAGAGAGCGCGCTGCTCTCTGCTGTC 13237
Qy 2487 CCTGCTGTCGACAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2546
Db 13238 CCTGCTGTCGACAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 13297
Qy 2547 AAAA 2550
Db 13298 AAAA 13301
```

```
RESULT 5
US-10-660-763-3
; Sequence 3, Application US/10660763
; Publication No. US20040063130A1
; GENERAL INFORMATION:
```

```
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001183DIVI
; CURRENT APPLICATION NUMBER: US/10/660,763
; CURRENT FILING DATE: 2003-09-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-660-763-3
```

```
Query Match 13.6%; Score 362.4; DB 17; Length 15297;
Best Local Similarity 99.7%; Pred. No. 1.2e-46;
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 2187 GGGGGGCGTCTGCTCTGCTCCAGAGCTGTGTCTGATGCGGTGTTCAAGCTCATGAGCAG 2246
Db 12938 GGGGGGCGTCTGCTCTGCTCCAGAGCTGTGTCTGATGCGGTGTTCAAGCTCATGAGCAG 12997
Qy 2247 TGCTGGGCTATGAGGCTGGGCGAGGCGCCAGCTTACAGCAATACAGAGAGCTGCAG 2306
Db 12998 TGCTGGGCTATGAGGCTGGGCGAGGCGCCAGCTTACAGCAATACAGAGAGCTGCAG 13057
Qy 2307 AGCATCCGAAAGCGGATCGGTGAGGCTGGGACCCCTTCTCAAGCTGTGAGCTCTGCA 2366
Db 13058 AGCATCCGAAAGCGGATCGGTGAGGCTGGGACCCCTTCTCAAGCTGTGAGCTCTGCA 13117
Qy 2367 GGGCTTGGTGGAGCTCTCTCAAGCGGCTTCAGACTATATGCTGACAGCTTTACAGTCTG 2426
Db 13118 GGGCTTGGTGGAGCTCTCTCAAGCGGCTTCAGACTATATGCTGACAGCTTTACAGTCTG 13177
Qy 2427 GACTCTGACACAGATCCACATGCGGCGAGAGAGAGCGCGCTGCTCTCTGCTGTC 2486
Db 13178 GACTCTGACACAGATCCACATGCGGCGAGAGAGAGCGCGCTGCTCTCTGCTGTC 13237
Qy 2487 CCTGCTGTCGACAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2546
Db 13238 CCTGCTGTCGACAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 13297
Qy 2547 AAAA 2550
Db 13298 AAAA 13301
```

```
RESULT 6
US-09-918-995-1503
; Sequence 1503, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1503
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(449)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-1503
```

Query Match 12.2%; Score 326.8; DB 10; Length 449;
Best Local Similarity 99.4%; Pred. No. 2.2e-41;
Matches 338; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1778 CGAGGGGGGCGATTCTTCTGACCTTCTCCGCAAGAGGGGGCCCGCTCGGGTGAAGAC 1837
120 GCGAGGGGGGCGATTCTTCTGACCTTCTCCGCAAGAGGGGGCCCGCTCGGGTGAAGAC 179

1838 TCTGTGCGATGATGATGGGGATGACAGCTGTGGCATGAGTACTTGAGAGCAAGTGTCTG 1897
180 TCTGTGCGATGATGATGGGGATGACAGCTGTGGCATGAGTACTTGAGAGCAAGTGTCTG 239

1898 CATCCACCGGGGACCTGGCTGTGCTGGAACCTGGTGAAGAGAGAAATGCTCTGAAGAT 1957
240 CATCCACCGGGGACCTGGCTGTGCTGGAACCTGGTGAAGAGAGAAATGCTCTGAAGAT 299

1958 CAGTGACTTTGGGATTTCCCGAGAGAGAACCGATGGGGTCTATGACAGCTTCAAGGGGCTT 2017
300 CAGTGACTTTGGGATTTCCCGAGAGAGAACCGATGGGGTCTATGACAGCTTCAAGGGGCTT 359

2018 CAGACAGTCTCCCGTGAAGTGAAGCGCACCTGAGCCCTTAACCTACGCGCTACTCTTC 2077
360 CAGACAGTCTCCCGTGAAGTGAAGCGCACCTGAGCCCTTAACCTACGCGCTACTCTTC 419

2078 CGAAGCGACGTGTGAGCTTTGGCATCTT 2107
420 CGAAGCGACGTGTGAGCTTTGGCATCTT 449

RESULT 7
US-09-948-802-5
; Sequence 5, Application US/09948802
; Publication No. US20020172981A1
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: NMI-090
; CURRENT APPLICATION NUMBER: US/09/948,802
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/387,212
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-948-802-5

Query Match 11.1%; Score 296.2; DB 9; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.1e-36;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

1684 CGAGATCTTGAAGCATGACGCAACCCCAACATCTGCTCATTTGTTGTGACACC 1743
20 CNAGATCTTGAAGCATGACGCAACCCCAACATCTGCTCATTTGTTGTGACACC 79

1744 AGAAGAGCCCATCTACATCTGATGAGCTTGTGAGGGGGGCGACTTCTGACCTTCC 1803
80 AGAAGAGCCCATCTACATCTGATGAGCTTGTGAGGGGGGCGACTTCTGACCTTCC 139

1804 TCCGCAAGAGGGGGGCGCTGCGGGTGAAGACTCTGCTGAGATGATGGGGATGACAG 1863
140 TCCGCAAGAGGGGGGCGCTGCGGGTGAAGACTCTGCTGAGATGATGGGGATGACAG 199

1864 CTGTGCGCATGAGTACTTGAGAGCAAGTGTGATCCACCGGGAAGCTTGGCTGTGGA 1923
200 CTGTGCGCATGAGTACTTGAGAGCAAGTGTGATCCACCGGGAAGCTTGGCTGTGGA 259

1924 ACTGCTGTGTGACAGAGAAATGTCTGTAAG-ATCAGTGAATCTT-GGATGTCCGAGA 1981

Db 260 ACTGCTGTGTGACAGAGAAATGTCTGTAAGATCAGTGAATCTTGGGATGTCCGAGA 319

QY 1982 GGAAGCC---GATGGGGTCTATGACAGCTTCAAGGGGGCTTCAAG 2020
Db 320 GGAAGCCGATTTGGGGGTCTATGACAGCTTCAAGGGGGCTTCAAG 361

RESULT 8
US-10-121-925-5
; Sequence 5, Application US/10121925
; Publication No. US20030104505A1
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: NMI-090
; CURRENT APPLICATION NUMBER: US/10/121,925
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US/09/948,802
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 09/387,212
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-10-121-925-5

Query Match 11.1%; Score 296.2; DB 15; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.1e-36;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

1684 CGAGATCTTGAAGCATGACGCAACCCCAACATCTGCTCATTTGTTGTGACACC 1743
20 CNAGATCTTGAAGCATGACGCAACCCCAACATCTGCTCATTTGTTGTGACACC 79

1744 AGAAGAGCCCATCTACATCTGATGAGCTTGTGAGGGGGGCGACTTCTGACCTTCC 1803
80 AGAAGAGCCCATCTACATCTGATGAGCTTGTGAGGGGGGCGACTTCTGACCTTCC 139

1804 TCCGCAAGAGGGGGGCGCTGCGGGTGAAGACTCTGCTGAGATGATGGGGATGACAG 1863
140 TCCGCAAGAGGGGGGCGCTGCGGGTGAAGACTCTGCTGAGATGATGGGGATGACAG 199

1864 CTGTGCGCATGAGTACTTGAGAGCAAGTGTGATCCACCGGGAAGCTTGGCTGTGGA 1923
200 CTGTGCGCATGAGTACTTGAGAGCAAGTGTGATCCACCGGGAAGCTTGGCTGTGGA 259

1924 ACTGCTGTGTGACAGAGAAATGTCTGTAAG-ATCAGTGAATCTT-GGATGTCCGAGA 1981
260 ACTGCTGTGTGACAGAGAAATGTCTGTAAGATCAGTGAATCTTGGGATGTCCGAGA 319

QY 1982 GGAAGCC---GATGGGGTCTATGACAGCTTCAAGGGGGCTTCAAG 2020
Db 320 GGAAGCCGATTTGGGGGTCTATGACAGCTTCAAGGGGGCTTCAAG 361

RESULT 9
US-10-280-576-19
; Sequence 19, Application US/10280576
; Publication No. US20040044405A1
; GENERAL INFORMATION:
; APPLICANT: Wolff, Matthew R.
; TITLE OF INVENTION: VASCULAR STENT OR GRAFT COATED OR IMPREGNATED WITH PROTEIN
; FILE REFERENCE: 09820.189
; CURRENT APPLICATION NUMBER: US/10/280,576
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 60/343,732

QY 2173 AGTTGTGAGAGAGGGGGGCGCTGCTGCCCCAGAGCTGTCTCTGTATGCGGTTC 2232
DB 2629 AGAGCATTAAGATGGGTACCGGTTGCCCCCTCTGTGACCTGCGCCCTGTATG 2688
QY 2233 GGCTCATGAGAGAGTGTGGGCTTATGAGCTGTGGGAGCGGCCCACTTC 2282
DB 2689 AGCTCATGAAGAACTGCTGGGATATGACCGTCCCGCGGACCACTTC 2738

RESULT 11
US-10-384-339C-1
Sequence 1, Application US/10384339C
Publication No. US20040175703A1
GENERAL INFORMATION:
APPLICANT: Kreutzer, Roland
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING EXPRESSION OF A TARGET GE
FILE REFERENCE: 20200/2002
CURRENT APPLICATION NUMBER: US/10/384,339C
PRIOR FILING DATE: 2003-03-07
PRIOR APPLICATION NUMBER: PCT/EP02/00152
PRIOR FILING DATE: 2002-01-09
PRIOR APPLICATION NUMBER: DE 10100586.5
PRIOR FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: DE 10155280.7
PRIOR FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: DE 1015841.3
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: DE 10160151.4
PRIOR FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 173
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2955
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
TITLE: Bph A1
PATENT DOCUMENT NUMBER: NM00532
US-10-384-339C-1

Query Match 7.4%; Score 198; DB 18; Length 2955;
Best Local Similarity 54.7%; Pred. No. 26-21;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1522 GGGTGTGAACCATGAGACCTGTGTGGGTGAGAGATTGACCGGGGAACCTTTGGCG 1581
DB 1874 GGGAGCTTATCCAGCGTGGCTGATGTGTGACACTGTCAATAGAGAAAGAGATTGGGG 1933
QY 1582 AAGTGTTCAGCGGACCGCTGCGAGCC-----GACAAACCCCTGTGTGGCGGTGAAT 1632
DB 1934 AAGTGTATGAGAGGAGCCCTCAGGCTCCAGCCAGACAGCTGTGAGACTGTGGCCATTAA 1993
QY 1633 CTGTGTGAGAGACGCTCCACCTGACCTGAAGGCCAAAGTTTCAAGAGAGGAGATCC 1692
DB 1994 CCTTAAAGACATATCCAGGTGGAGAGTGTGAACTTCTTCAAGAGGCACTATGA 2053
QY 1693 TGAAGAGTACAGCCACCCCAACATGTGTCTCATTTGGTGTCTGACCCAGAAACAGC 1752
DB 2054 TGGGCGCAGTTTACGACCCGCAATTTCTGATCTGAAAGCGTCGCAAAAGCGAAAGC 2113
QY 1753 CCATCTACATCGTATGAGAGCTTGTGACGGGGGCGACTTCTGACCTTCTCCGACGG 1812
DB 2114 CCATCTATGATCATCAAGAAATTTATGAGAAATGCAACCCCTGATGCTTCTGAGGAGC 2173
QY 1813 AGGGGGCCGGCCCTGGGGGGTGAAGACTGTGTGAGATGAGTGGGGGAGATGACGCTGGCA 1872
DB 2174 GGGAGACCAAGCTGTTCCTGGGCACTTGTGCGCATGTGCAAGGCAATGACTTGGCA 2233
QY 1873 TGAAGTACCTGAGAGCAAGTGTGTGACCCGAGACCTGTGTGTGCGAACTGTGCTGG 1932
DB 2234 TGAAGTACCTGAGTATCAACAATTAATGTGCAACCGGAGCCGTGGTGCAGAAACATCTTGG 2293

QY 1933 TGACAGAGAAAGATGCTTGAATCATGATGACTTTGGATGTCCCGAGAGAAACCGATG 1992
DB 2294 TGAATCAAAACCTGTGCTGCAAGGTGTGACTTTTGGCTGTAGCTGCTTCTGATGACT 2353
QY 1993 GGGTCTATGACCTTCAGGGGGGCTTCAGAAAGTCCCGCTGAAGTGAACCGACCTGAG 2052
DB 2354 TTGATGGCACATATCAAGAAACCGAGGAGAAAGATCCCTATCCGTTGGAAGCCCTGAA 2413
QY 2053 CCTTAACTAAGCGCGCTACTCTCCGAAAGGAGAGTGTGAGCTTTGGCATTTCTCT 2112
DB 2414 CCATTGCCATCGATGATTTCAACACAGCCAGGATGTGTGAGCTTTGGAAATGTATGT 2473
QY 2113 GGGAGACCTTCAGCCCTGGGGGCGCTCCCTATGCCAAGCTTCAGCATGACAGACGCG 2172
DB 2474 GGGAGTGTGAGCTTTGGGGACAAAGCTTATGAGGAGATGACATTCAGAGGTTATGA 2533
QY 2173 AGTTGTGAGAGAGGGGGGCGCTGCTGCCCAAGCTGTGTCTGTATGCGGTTC 2232
DB 2534 AGAGCATTAAGATGGGTATCGGGTTGCCCTCTGTGTGACGTGCCCTGCCCTGTATG 2593
QY 2233 GGCTCATGAGAGAGTGTGGGCTTATGAGCTGTGGGAGCGGCCCACTTC 2282
DB 2594 AGCTCATGAAGAACTGCTGGGATATGACCGTCCCGCGGACCACTTC 2643

RESULT 12
US-09-967-768A-144
Sequence 144, Application US/09967768A
Patent No. US20020150877A1
GENERAL INFORMATION:
APPLICANT: Augustus, Meena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-72
CURRENT APPLICATION NUMBER: US/09/967,768A
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/60/236,109
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,034
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,111
PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 325
SOFTWARE: PatentIn version 3.0
SEQ ID NO 144
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
US-09-967-768A-144

Query Match 7.4%; Score 198; DB 9; Length 3370;
Best Local Similarity 54.7%; Pred. No. 26-21;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1522 GGGTGTGAACCATGAGACCTGTGTGGGTGAGAGATTGACCGGGGAACCTTTGGCG 1581
DB 1967 GGGAGCTTATCCAGCGTGGCTGATGTGTGACACTGTCAATAGAGAAAGAGATTGGGG 2026
QY 1582 AAGTGTTCAGCGGACCGCTGCGAGCC-----GACAAACCCCTGTGTGGCGGTGAAT 1632
DB 2027 AAGTGTATGAGAGGAGCCCTCAGGCTCCAGCCAGACAGCTGTGAGACTGTGGCCATTAA 2086
QY 1633 CTGTGTGAGAGACGCTCCACCTGACCTGAAGGCCAAAGTTTCAAGAGAGGAGATCC 1692
DB 2087 CCTTAAAGACATATCCAGGTGGAGAGTGTGAACTTCTTCAAGAGGCACTATGA 2146
QY 1693 TGAAGAGTACAGCCACCCCAACATGTGTCTCATTTGGTGTCTGACCCAGAAACAGC 1752
DB 2147 TGGGCGCAGTTTACGACCCGCAATTTCTGATCTGAAAGCGTCGCAAAAGCGAAAGC 2206
QY 1753 CCATCTACATCGTATGAGAGCTTGTGACGGGGGCGACTTCTGACCTTCTCCGACGG 1812
DB 2207 CCATCTATGATCATCAAGAAATTTATGAGAAATGCAACCCCTGATGCTTCTGAGGAGC 2266

QY 1813 AGGGGCGCGCTGCGGGTGAAGACTCTGCTGCAGATGCGGGGAGTGAAGTGTGGA 1872
DB 2267 GGGAGAGACCACTGCTGCTGCGGACCTGAGTGGCCATGCTGAGGGGCAATGACTGGCA 2326
QY 1873 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGAGCTGCTGCTGGAATGCTGG 1932
DB 2327 TGAATACCTGAGTATCAATATATGTCACCGGAGCTGCTGCTGCAAGAACATCTTG 2386
QY 1933 TGAAGAGAAATGCTGGAAGATGATGATCTTGGATGTCCTGAGAGAACCGGATG 1992
DB 2387 TGAATCAAACTGTGCTGCAAGTGTGATCTTGGATGTCCTGAGTACT 2446
QY 1993 GGGTATGAGCGCTGAGGGGCGCTGAGCAAGTGTGAGTGTGAGTGTGAGTGTGAGT 2052
DB 2447 TGAATGAGCAATGAGAAACCGAGGAGGAAAGTGTGATCTTGGATGTCCTGAGT 2506
QY 2053 CCTTAATGAGCGCGCTGCTGCTGCGAAAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2112
DB 2507 CCATTGCCATGAGTCTTACACGAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2566
QY 2113 GGGAGCTTTCAGCTGCGGGGCTCCCTGATCCCACTGAGTGTGAGTGTGAGTGTGAGT 2172
DB 2567 GGGAGGTGTGAGCTTGGGGGAGAGCTTATGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2626
QY 2173 AGTTGTGAGAGAGGGGGGCGCTGCTGCGGAGCTGTGAGTGTGAGTGTGAGTGTGAGT 2232
DB 2627 AGAGCATTAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2686
QY 2233 GGCTCATGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2282
DB 2687 AGCTCATGAAGAACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2736

RESULT 13

US-10-354-358-101
Sequence 101, Application US/10354358
Publication No. US20030157082A1

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Hunter, John Joseph
APPLICANT: Macheth, Kyle J.
APPLICANT: Tsai, Fong-Ying
APPLICANT: Lesoon, Andrea
APPLICANT: Lightcap, Eric S.
APPLICANT: Williamson, Mark A.
APPLICANT: Rudolph-Owen, Laura A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
TITLE OF INVENTION: 9252, 9389, 1642, 85268, 10297, 1588, 9525, 14124, 4469,
TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
TITLE OF INVENTION: 16534, 68862, 9011, 14031, 6178, 21225, 1420, 32336, 2099,
TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
FILE REFERENCE: MPIO2-020P1RNM0NIM
CURRENT APPLICATION NUMBER: US/10/354,358
CURRENT FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/385,023

PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 101
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (94)...(3048)
US-10-354-358-101

Query Match 7.4%; Score 198; DB 16; Length 3370;
Best Local Similarity 54.7%; Pred. No. 28-21;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1522 GGGTCTGACCATGAGACCTGTGTTGGGTGAGCAGATTGAGCGGGAACTTTGGG 1581
DB 1967 GGGAGCTTGTATCAGCGCTGCTGATGTGTGAGCACTGTCTATGAGAGAGAGTTGGG 2026
QY 1582 AGTGTTCAGCGAGCGCTGCGAGCC-----GACAAACCTGTGTGGCGTGAAGT 1632
DB 2027 AGGTGTATGAGAGAGACCTTCAAGGCTCCAGCAGAGCTGCAAGACTGTGCTTAA 2086
QY 1633 CTGTGTGAGAGAGCTCTCCACCTGACTTCAGAGCCAGTTTTCAGAGAGAGAGATCC 1692
DB 2087 CCTTAAAGACACATCCAGAGTGGCCAGTGTGTGAACTTCTTGAAGAGAGAACTATCA 2146
QY 1693 TGAACAGTACAGCCAGCCCAATCTGCGCTCATTTGATGTGTGACCCAGAGAGAGC 1752
DB 2147 TGGGCGAGTTTACCAACCGCATATTTGTGATGTGAGAGGCGTCTGCAAGAGAGAGC 2206
QY 1753 CCATCTACATCGTCAATGAGCTTGTGAGGGGGGAGTCTTCTGCACTTCTCCGACCG 1812
DB 2207 CATCATGATCATCAAGAAATTTATGAGAAATGAGAGCCCTGATGCTTCTGAGAGAGC 2266
QY 1813 AGGGGCGCGCTGCGGGTGAAGACTGTGTCAGATGTGTGAGAGATGAGAGTGTGAGCA 1872
DB 2267 GGGAGGAGCAGCTGTGCTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2326
QY 1873 TGAAGTACCTGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932
DB 2327 TGAAGTACCTGAGTATCAATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2386
QY 1933 TGAAGAGAAATGCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
DB 2387 TGAATCAAACTGTGCTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2446
QY 1993 GGGTCTGAGAGCTTCAAGGGGCTTCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2052
DB 2447 TGAATGAGCAATACAAACCGAGAGGAAATGCTTCTGATGATGATGATGATGATGATGATGATGATGAT 2506
QY 2053 CCTTAATGAGCGCGCTGCTGCTGCGAAAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2112
DB 2507 CCATTGCCATGAGTCTTACACGAGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2566
QY 2113 GGGAGCTTTCAGCTGCGGGGCTCCCTGATCCCACTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2172
DB 2567 GGGAGGTGTGAGCTTGGGGAGAGAGCTTATGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2626
QY 2173 AGTTGTGAGAGAGGGGGGCGTGTGCTGCGCAAGTGTGTGATGATGATGATGATGATGATGATGAT 2232
DB 2627 AGAGCATTAAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2686
QY 2233 GGCTCATGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2282
DB 2687 AGCTCATGAAGAACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2736

RESULT 14
US-10-210-120-19
Sequence 19, Application US/10210120
Publication No. US2003015736A1
GENERAL INFORMATION:
APPLICANT: Chinnaiyan, Arul M.
APPLICANT: Rubin, Mark A.
APPLICANT: Steekumar, Arun
TITLE OF INVENTION: Expression Profile of Prostate Cancer
FILE REFERENCE: UM-07221
CURRENT APPLICATION NUMBER: US/10/210,120
CURRENT FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: US 60/309,581
PRIOR FILING DATE: 2001-08-02
PRIOR APPLICATION NUMBER: US 60/334,468
PRIOR FILING DATE: 2001-11-15
NUMBER OF SEQ ID NOS: 123
SOFTWARE: PatentIn version 3.2
SEQ ID NO: 19
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
US-10-210-120-19

Query Match 7.4%; Score 198; DB 16; Length 3370;

Best Local Similarity 54.7%; Pred. No. 2e-21;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

1522 GGGTGTGAACCATGAGACCTGTGTGGGTGAGACGATTGAGACGGGGACCTTTGGCG 1581
1967 GGGAGCTTGAATCCAGGTGCTATGTGACACTGTCAATGAGAGAGAGATTGGGG 2026
1582 AAGTGTTCAGCGGACGCTGCGAGCC-----GACAAACCTGTGGCGGTGAAGT 1632
2027 AAGTGTATGAGGAGACCTTCAGGCTCCCGACGAGACCTGCAAGACTGTGGCATTAAGA 2086
1633 CTGTGTGAGAGAGAGCTTCCCACTGACTCTCAAGGCGCAAGTTTCAACAGAGAGATCC 1692
2087 CCTTAAAGACATATCCCGAGGTGGCCAGTGTGGAACCTTCTTGAAGGCAACTATCA 2146
1693 TGAAGAGTACAGACCCCAACATCATGTGGGTCTCAATTGGTGTCTGACCCAGAGAGAGC 1752
2147 TGGGCCAGTTTACGACCCGCAATTTCTGATCTGAGAGGCGTGTCAAAAGCGAAAGC 2206
1753 CCATTCATATGTCATGAGAGCTTGTGAGGGGGCGAATTCTGACCTTCTCCGACGAG 1812
2207 CGATTCATGATCATCAGAGATTTATGAGAGATGCAAGCCGTGATGCTTCTGAGGGAGC 2266
1813 AGGGGGCCGCTGCGGGTGAAGACTGTCTGAGATGTGGGGATGACGCTGTGGCA 1872
2267 GGGAGGACCAAGCTGTGCTCCCTGGGAGTGGCCATGTGCAAGGGCATAGCATCTGGCA 2326
1873 TGAAGTACCTGAGAGCAAGTGTGATCCACGGGAGCTGGGTGCTCGGAATGCTGTGG 1932
2327 TGAATCACTTCAATATCAATATATGTCCACGGGAGCTGGGTGCGAGAAACATTTGG 2386
1933 TGAAGAGAGAGATGTCTGAGAGATGATGATCTTTGGGATGTCCGAGAGAGAGCCGATG 1992
2387 TGAATCAAAACCTGTGTGTGAGAGTGTCTGACTTTGGCCGATCTCGGCTCTGAGATGAC 2446
1993 GGGTCTATGAGAGCTTCAAGGGGGCTTCAAGCAAGTCCCGTGAATGAGACCGCACTGAGG 2052
2447 TTGATGAGCATATCGAAACCCAGAGAGAGAAAGATCCCTATCCGTGAGACGCCCCCTGAG 2506
2053 CCCTTAACATCGGCGCTACTCTCCGAAAGGAGCGATGAGAGCTTTGGCATCTGTGCT 2112
2507 CCATTTGCCATCGGATCTTTCACACAGCCAGCATGTGTGAGAGCTTTGGATTTGATGT 2566
2113 GGGAGACCTTTCAGAGCTGGGGGCTCCCTATCCCAACTCAGCAATCAGCAGACAGCGG 2172
2567 GGGAGGTGCTGAGCTTTGGGGAGACAGGCTTATGGGGAGATGACCAATCAGAGAGTTATGA 2626

Qy 2173 AGTTGTGAGAGAGGGGGCGGCTGCTGCCAGAGCGGTGTCTGATGCCGTGTCA 2232
Db 2627 AGAGCATTAAGATGAGGTACCGGTTGCCCTCTCTGTGACCTGCCCTCTGTATG 2686
Qy 2233 GGCTCATGAGAGAGTGTGCTGAGCTTATGAGCTTGGAGCGGCCAGCTTC 2282
Db 2687 AGCTCATGAAGAACTGCTGGGATATGACCGTGTCCCGGGCCACACTTC 2736

RESULT 15
US-10-776-827-82

Sequence 82, Application US/10776827
Publication No. US20040132086A1

GENERAL INFORMATION:

APPLICANT: Horwitz, Kathryn

APPLICANT: Richer, Jennifer

TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relate

FILE REFERENCE: 2848-39

CURRENT APPLICATION NUMBER: US/10/776,827

CURRENT FILING DATE: 2004-02-10

PRIOR APPLICATION NUMBER: US/09/814,915

PRIOR FILING DATE: 2002-03-21

PRIOR APPLICATION NUMBER: 60/214,870

PRIOR FILING DATE: 2000-06-28

NUMBER OF SEQ ID NOS: 108

SOFTWARE: PatentIn version 3.1

SEQ ID NO: 82

LENGTH: 3370

TYPE: DNA

ORGANISM: Homo sapiens

US-10-776-827-82

Query Match 7.4%; Score 198; DB 18; Length 3370;

Best Local Similarity 54.7%; Pred. No. 2e-21;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

1522 GGGTGTGAACCATGAGACCTGTGTGGGTGAGACGATTGAGACGGGGACCTTTGGCG 1581
1967 GGGAGCTTGAATCCAGGTGCTATGTGACACTGTCAATGAGAGAGAGATTGGGG 2026
1582 AAGTGTTCAGCGGACGCTGCGAGCC-----GACAAACCTGTGGCGGTGAAGT 1632
2027 AAGTGTATGAGGAGACCTTCAGGCTCCCGACGAGACCTGCAAGACTGTGGCATTAAGA 2086
1633 CTGTGTGAGAGAGAGCTTCCCACTGACTCTCAAGGCGCAAGTTTCAACAGAGAGATCC 1692
2087 CCTTAAAGACATATCCCGAGGTGGCCAGTGTGGAACCTTCTTGAAGGCAACTATCA 2146
1693 TGAAGAGTACAGACCCCAACATCATGTGGGTCTCAATTGGTGTCTGACCCAGAGAGAGC 1752
2147 TGGGCCAGTTTACGACCCGCAATTTCTGATCTGAGAGGCGTGTCAAAAGCGAAAGC 2206
1753 CCATTCATATGTCATGAGAGCTTGTGAGGGGGCGAATTCTGACCTTCTCCGACGAG 1812
2207 CGATTCATGATCATCAGAGATTTATGAGAGATGCAAGCCGTGATGCTTCTGAGGGAGC 2266
1813 AGGGGGCCGCTGCGGGTGAAGACTGTCTGAGATGTGGGGATGACGCTGTGGCA 1872
2267 GGGAGGACCAAGCTGTGCTCCCTGGGAGTGGCCATGTGCAAGGGCATAGCATCTGGCA 2326
1873 TGAAGTACCTGAGAGCAAGTGTGATCCACGGGAGCTGGGTGCTCGGAATGCTGTGG 1932
2327 TGAATCACTTCAATATCAATATATGTCCACGGGAGCTGGGTGCGAGAAACATTTGG 2386
1933 TGAAGAGAGAGATGTCTGAGAGATGATGATCTTTGGGATGTCCGAGAGAGAGCCGATG 1992
2387 TGAATCAAAACCTGTGTGTGAGAGTGTCTGACTTTGGCCGATCTCGGCTCTGAGATGAC 2446
1993 GGGTCTATGAGAGCTTCAAGGGGGCTTCAAGCAAGTCCCGTGAATGAGACCGCACTGAGG 2052
2447 TTGATGAGCATATCGAAACCCAGAGAGAGAAAGATCCCTATCCGTGAGACGCCCCCTGAG 2506

```
Qy 2053 CCTTAAGGCGCGCTACTCCTCGAAGGAGCGTGTGAGCTTTGGCATCTTGTCT 2112
    |||||
Db 2507 CCATTGCCCATGGATCTTCAACACAGCCAGCGATGTGTGAGCTTTGGATTGTATGT 2566
    |||||
Qy 2113 GGGAGACTTCAAGCTGGGGGCTCCCTTATCCAACTCAGCAATCAGCAGACAGGG 2172
    |||||
Db 2567 GGGAGGCTGTGAGCTTTGGGACAGAGCTTATGGGAGATGAGCAATCAGAGGTTATGA 2626
    |||||
Qy 2173 AGTTGTGAGAGAGGGGGCGTCTGCGCTGCCAGAGCTGTCTCGATGCCGTGTCA 2232
    |||||
Db 2627 AGAGCATTTAGAGATGGGTACCGGTGCCCCCTCTGTGACTGCCCTTCTGTATG 2686
    |||||
Qy 2233 GGCTCATGAGAGCAAGTCTGGGCTTATGAGGCTTGGGAGCGGCCCAAGCTTC 2282
    |||||
Db 2687 AGCTCATGAAGAACTGTGGGCAATGACCGTGCCCGGCCCAACTTC 2736
    |||||
```

Search completed: March 19, 2005, 19:55:48
Job time : 1398.78 secs

Db 361 GGGCCCTGAGCAAGTGAAGCTGATCCGGGAAAGGAGAGCTTCCGAAGACTTAA 420
QY 421 GCGAGAGTGGAGAGCTGACAGAGAGCTCAACAAAGCCCAAGCCAGAGATTGAA 480
Db 422 GCGAGAGTGGAGAGCTGACAGAGAGCTCAACAAAGCCCAAGCCAGAGATTGAA 480
QY 481 AGCTGAAGAGCCAGTACCGAGCTCTGGCAAGGAGCAAGTCCAAAGCCAAAGCAATACC 540
Db 481 AGCTGAAGAGCCAGTACCGAGCTCTGGCAAGGAGCAAGTCCAAAGCCAAAGCAATACC 540
QY 541 AGAGAGCCAGCAAGAAAGCAAGAGCCGTGACAAAGCCAAAGCAAGTATGTCGAGCTGT 600
Db 541 AGAGAGCCAGCAAGAAAGCAAGAGCCGTGACAAAGCCAAAGCAAGTATGTCGAGCTGT 600
QY 601 GGAAGCTCTTGTGTCACCAACCGCTATGTGTCGGGCGTGGGCTGGGAGCTTACAC 660
Db 601 GGAAGCTCTTGTGTCACCAACCGCTATGTGTCGGGCGTGGGCTGGGAGCTTACAC 660
QY 661 ACCAGAGCCAGCAAGAGCTCTGCGCCGAGCCGTGCGAGTCACTGAGAGACTGACAG 720
Db 661 ACCAGAGCCAGCAAGAGCTCTGCGCCGAGCCGTGCGAGTCACTGAGAGACTGACAG 720
QY 721 AGAGAGTGGCTTGCATCTTGAAGAGAAATCTGAGAGAAATCTGAGATTAGCAGCTGG 780
Db 721 AGAGAGTGGCTTGCATCTTGAAGAGAAATCTGAGAGAAATCTGAGATTAGCAGCTGG 780
QY 781 TGAAGATAGAGTGGTGGCCATTCAACGGGAGATGGCTGAGAGTGTCCCGGATCCAG 840
Db 781 TGAAGATAGAGTGGTGGCCATTCAACGGGAGATGGCTGAGAGTGTCCCGGATCCAG 840
QY 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGATAGGTCGAGCACTGACCTCCACCT 900
Db 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGATAGGTCGAGCACTGACCTCCACCT 900
QY 901 GTGTCAAGTTCAGTCACTGCTTGAAGAGGAGTGAACCGCTGAGAGCTGGAGCTCC 960
Db 901 GTGTCAAGTTCAGTCACTGCTTGAAGAGGAGTGAACCGCTGAGAGCTGGAGCTCC 960
QY 961 AGCTGAACGAGCTGAGCTGAGAGAGGTCGAGCAACGCTGACCTCAAGTGAAGAG 1020
Db 961 AGCTGAACGAGCTGAGCTGAGAGAGGTCGAGCAACGCTGACCTCAAGTGAAGAG 1020
QY 1021 TGGCTGTGCGACAGAGTGTGTCAGGCGGAGAGAGTGTGAAGCAAGCTTCAACAG 1080
Db 1021 TGGCTGTGCGACAGAGTGTGTCAGGCGGAGAGAGTGTGAAGCAAGCTTCAACAG 1080
QY 1081 AGCTCGGAAATGAAGAGAGAAACCCACCCCGGAGCGGTCGAGCTTGGGCAAGA 1140
Db 1081 AGCTCGGAAATGAAGAGAGAAACCCACCCCGGAGCGGTCGAGCTTGGGCAAGA 1140
QY 1141 GCGAAGTGTGCAAGAGCACTGAGGAGTGAAGAGTGTGTCGAGAGCCAGAGC 1200
Db 1141 GCGAAGTGTGCAAGAGCACTGAGGAGTGAAGAGTGTGTCGAGAGCCAGAGC 1200
QY 1201 TGAAGGCGCAGAGAGTGTGTCAGACAGCTGAGAGCACTGGGCGCGGAGAGCC 1260
Db 1201 TGAAGGCGCAGAGAGTGTGTCAGACAGCTGAGAGCACTGGGCGCGGAGAGCC 1260
QY 1261 CGCTGTGTGCTCTGTCAGAGATGACCGCACTTCCCTGTCGAGAGAGAGAG 1320
Db 1261 CGCTGTGTGCTCTGTCAGAGATGACCGCACTTCCCTGTCGAGAGAGAGAG 1320
QY 1321 AGGGGGGAGAGAGCAACGCTGAGAGTCTTGAAGAGCCATCTCAAGAAATCTTCCGC 1380
Db 1321 AGGGGGGAGAGAGCAACGCTGAGAGTCTTGAAGAGCCATCTCAAGAAATCTTCCGC 1380
QY 1381 CCAAGTCTGCAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 CCAAGTCTGCAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 TGAAGCACTTGAAG 1500
Db 1441 TGAAGCACTTGAAG 1500

QY 1501 GGGCTGTGCGCAAGAGCAAGTGGTGTGAACCATGAGAGAGCTGTGTGGTGAAGAG 1560
Db 1501 GGGCTGTGCGCAAGAGCAAGTGGTGTGAACCATGAGAGAGCTGTGTGGTGAAGAG 1560
QY 1561 TTGAGAGGGGAGAACTTTGGCGAAAGTTCAGCGGACGCTGCGAGAGCCAAACCTTGG 1620
Db 1561 TTGAGAGGGGAGAACTTTGGCGAAAGTTCAGCGGAGAGCTGCGAGAGCCAAACCTTGG 1620
QY 1621 TGGCGGTGAAGCTTGTGAGAGAGAGCTCCCACTGACCTCAAGGCAAGTTCATCAAG 1680
Db 1621 TGGCGGTGAAGCTTGTGAGAGAGAGCTCCCACTGACCTCAAGGCAAGTTCATCAAG 1680
QY 1681 AAGCGAGATCTGAAAGAGTACAGAGCAACCCCAACATGTCGCTCATTTGGTGTGCA 1740
Db 1681 AAGCGAGATCTGAAAGAGTACAGAGCAACCCCAACATGTCGCTCATTTGGTGTGCA 1740
QY 1741 CCGAAGAGAGCCCAATCTATCATGAGCTTGTGAGAGGAGGAGAGAGAGAGAGAG 1800
Db 1741 CCGAAGAGAGCCCAATCTATCATGAGCTTGTGAGAGGAGGAGAGAGAGAGAGAG 1800
QY 1801 TCTTCGAG 1860
Db 1801 TCTTCGAG 1860
QY 1861 CAGCTGTGAGAGTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1861 CAGCTGTGAGAGTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1921 GGAAGCTGTGAG 1980
Db 1921 GGAAGCTGTGAG 1980
QY 1981 AGGAAGCCAG 2040
Db 1981 AGGAAGCCAG 2040
QY 2041 CCGAGCTGAG 2100
Db 2041 CCGAGCTGAG 2100
QY 2101 GCAATCTGTGAG 2160
Db 2101 GCAATCTGTGAG 2160
QY 2161 AGCAG 2220
Db 2161 AGCAG 2220
QY 2221 ATGCGGTGTCAGAGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
Db 2221 ATGCGGTGTCAGAGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
QY 2281 TGAAGCACTTGAAG 2340
Db 2281 TGAAGCACTTGAAG 2340
QY 2341 CCGTTCAG 2400
Db 2341 CCGTTCAG 2400
QY 2401 TATGCTGAG 2460
Db 2401 TATGCTGAG 2460
QY 2461 ATGAG 2520
Db 2461 ATGAG 2520
QY 2521 AACATTAAG 2580
Db 2521 AACATTAAG 2580


```

Db 1681 AAGCGAGATCTCGAAGCAGTACAGCACCCCAACATCGTGCCTCATTTGGTGTCTGCA 1740
Qy 1741 CCCAGAAAGCAGCCCATCTACATGTCTATGAGAGCTTTGCGAGGGGGGCGACTTCTGA 1800
Db 1741 CCCAGAAAGCAGCCCATCTACATGTCTATGAGAGCTTTGCGAGGGGGGCGACTTCTGA 1800
Qy 1801 TCCTCGCAGCGAGGGGGGCGCGCTGCGGGGTGAAGA CTCTGCTGCAAGATGTTGGGAGATG 1860
Db 1801 TCCTCGCAGCGAGGGGGGCGCGCTGCGGGGTGAAGA CTCTGCTGCAAGATGTTGGGAGATG 1860
Qy 1861 CAGCTCTGCGCATGAGTATCTGAGAGCAAGTGTCTGATCCACCGGAGCTTGCTGCTC 1920
Db 1861 CAGCTCTGCGCATGAGTATCTGAGAGCAAGTGTCTGATCCACCGGAGCTTGCTGCTC 1920
Qy 1921 GGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 1921 GGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Qy 1981 AGGAGCGCGATGAGGCTTATGAGAGCTGAGAGGCGCGCTGAGCAAGTCCCGTGAAGTGA 2040
Db 1981 AGGAGCGCGATGAGGCTTATGAGAGCTGAGAGGCGCGCTGAGCAAGTCCCGTGAAGTGA 2040
Qy 2041 CCGCAGCTGAGGCGCTTAACTAAGCGCGCTAATCTCCGAAAGCGAGCTGTGAGCTTTG 2100
Db 2041 CCGCAGCTGAGGCGCTTAACTAAGCGCGCTAATCTCCGAAAGCGAGCTGTGAGCTTTG 2100
Qy 2101 GCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 2101 GCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Qy 2161 AGGAGCAGCGGAGGCTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
Db 2161 AGGAGCAGCGGAGGCTTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
Qy 2221 ATGCGGTGTTCAAGCTCTATGAGAGCAGTGTGAGGCTTATGAGAGGAGGAGGAGGAGGAG 2280
Db 2221 ATGCGGTGTTCAAGCTCTATGAGAGCAGTGTGAGGCTTATGAGAGGAGGAGGAGGAGGAG 2280
Qy 2281 TCGAGCAGCATCTACAGAGAGCTGAGAGCATCCGAAAGCGGAGCTGAGGAGCTGAGGAG 2340
Db 2281 TCGAGCAGCATCTACAGAGAGCTGAGAGCATCCGAAAGCGGAGCTGAGGAGCTGAGGAG 2340
Qy 2341 CCTTCTCAAGCTGAGGAGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGTCA 2400
Db 2341 CCTTCTCAAGCTGAGGAGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGTCA 2400
Qy 2401 TATGCTGACAGCTCTTCAAGAGCTCTGAGATCTGCGCAGCAGCATCCACTGCGAGG 2460
Db 2401 TATGCTGACAGCTCTTCAAGAGCTCTGAGATCTGCGCAGCAGCATCCACTGCGAGG 2460
Qy 2461 ATGAGAGGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2520
Db 2461 ATGAGAGGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2520
Qy 2521 AACCAATTAACCACTTGTGCGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Db 2521 AACCAATTAACCACTTGTGCGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Qy 2581 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2640
Db 2581 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2640
Qy 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674
Db 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674

```

RESULT 3
 US-09-817-180-3
 ; Sequence 3, Application US/09817180
 ; Patent No. 6340584
 ; GENERAL INFORMATION:

```

; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183
; CURRENT FILING DATE: US/09/817,180
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Human
US-09-817-180-3

```

Query Match 13.6%; Score 362.4; DB 3; Length 15297;
 Best Local Similarity 99.7%; Pred. No. 5,1e-55;
 Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 2187 GGGGGGCGTGTGCGCTGCGCCAGAGCTGTGCTGATGCGGTGTTCAAGCTCATGAGAGAG 2246
Db 12938 GGGGGGCGTGTGCGCTGCGCCAGAGCTGTGCTGATGCGGTGTTCAAGCTCATGAGAGAG 12997
Qy 2247 TGTGTGCGCTATGAGCTGTGAGCGGCGCCAGCTTACAGCAGCATTAACAGAGAGTGCAG 2306
Db 12998 TGTGTGCGCTATGAGCTGTGAGCGGCGCCAGCTTACAGCAGCATTAACAGAGAGTGCAG 13057
Qy 2307 AGCATCCGAAAGCGGAGTGTGAGAGGCTGTGAGACCCCTTCTCAAGCTGTGCTGTGCA 2366
Db 13058 AGCATCCGAAAGCGGAGTGTGAGAGGCTGTGAGACCCCTTCTCAAGCTGTGCTGTGCA 13117
Qy 2367 GGGCTGAGTGCAGAGCTCTTCAAGGAGGCTCCAGCTCATATGCTGACAGCTTTCAAGTCTG 2426
Db 13118 GGGCTGAGTGCAGAGCTCTTCAAGGAGGCTCCAGCTCATATGCTGACAGCTTTCAAGTCTG 13177
Qy 2427 GACTCTGCGCACAGAGTCCAGAGCTGCGGAGAGATGACGCGGCTGTCTCTGTGTC 2486
Db 13178 GACTCTGCGCACAGAGTCCAGAGCTGCGGAGAGATGACGCGGCTGTCTCTGTGTC 13237
Qy 2487 CCTGTGCTGCGCACAGGCTTCTTCCGGGCGAGAAACAATAAACAATTGCGCCACTG 2546
Db 13238 CCTGTGCTGCGCACAGGCTTCTTCCGGGCGAGAAACAATAAACAATTGCGCCACTG 13297
Qy 2547 AAAA 2550
Db 13298 AACA 13301

```

```

RESULT 4
US-10-003-295-3
; Sequence 3, Application US/10003295
; Patent No. 6686187
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183D1V
; CURRENT APPLICATION NUMBER: US/10/003,295
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-3

```

Query Match 13.6%; Score 362.4; DB 4; Length 15297;
 Best Local Similarity 99.7%; Pred. No. 5,1e-55;
 Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2187 GGGGGGCGTGTGCGCTGCGCCAGAGCTGTGCTGATGCGGTGTTCAAGCTCATGAGAGAG 2246

Db 12938 GGGGGCCGCTGCTCCCTCCAGAGCTGTCTGATCCCTGTTCAAGGCTCATGAGCCAG 12997
Qy 2247 TCGTGGGCTTATGAGCTTGGGAGCGGCCAGCTTGACCACTATACAGAGAGCTGACG 2306
Db 12998 TGTGGGCTTATGAGCTTGGGAGCGGCCAGCTTGACCACTATACAGAGAGCTGACG 13057
Qy 2307 AGCATCCGAAAGGGGAGTGGTGAAGGCTGGGACCCCTTCTCAAGCTGTGGCTTGTGA 2366
Db 13058 AGCATCCGAAAGGGGAGTGGTGAAGGCTGGGACCCCTTCTCAAGCTGTGGCTTGTGA 13117
Qy 2367 GGGCTGAGTGAAGCTCTGAGCGGCTGAGGCTCATATGAGTGAAGCTTTCACAGTCTG 2426
Db 13118 GGGCTGAGTGAAGCTCTGAGCGGCTGAGGCTCATATGAGTGAAGCTTTCACAGTCTG 13177
Qy 2427 GACTCTGACCCAGCATCCACACTGCGGAGAGTGAAGCGCGTGTCTCTGTGTGTC 2486
Db 13178 GACTCTGACCCAGCATCCACACTGCGGAGAGTGAAGCGCGTGTCTCTGTGTGTC 13237
Qy 2487 CCGTCTGCTGCCAGGCGCTTCTTCTCGGGGAGAAAACATTAACCACTTGTGCCACTG 2546
Db 13238 CCGTCTGCTGCCAGGCGCTTCTTCTCGGGGAGAAAACATTAACCACTTGTGCCACTG 13297
Qy 2547 AAAA 2550
Db 13298 AACCA 13301

RESULT 5

US-09-387-212-5
Sequence 5, Application US/09387212A
Patent No. 6309849
GENERAL INFORMATION:
APPLICANT: ROBINSON, KEITH E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
FILE REFERENCE: NMT-090
CURRENT APPLICATION NUMBER: US/09/387, 212A
CURRENT FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentlin Ver. 2.0
SEQ ID NO 5
LENGTH: 361
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-387-212-5

Query Match 11.1%; Score 296.2; DB 3; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.1e-43;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

Qy 1684 CGAGATCTCGAAGCAGTACAGCCCAACATCGGTCTCATTTGATGTGACACC 1743
Db 20 CNAAGATCTCGAAGCAGTACAGCCCAACATCGGTCTCATTTGATGTGACACC 79
Qy 1744 AGAAGACCCCATCTATCATGTGAGAGCTTGTGAGGGGGGCACTTCTGACCTTCC 1803
Db 80 AGAAGACCCCATCTATCATGTGAGAGCTTGTGAGGGGGGCACTTCTGACCTTCC 139
Qy 1804 TCCGCAACGAGGGGGGCGGCTCGGGGTGAAGACTCTGTGCAAGATGTGGGAGATGAC 1863
Db 140 TCCGCAACGAGGGGGGCGGCTCGGGGTGAAGACTCTGTGCAAGATGTGGGAGATGAC 199
Qy 1864 CTGCTGCAATGAGTACTTGAAGAGCAAGTGTGATCCACCGGAGACTGTGGCTGTGGA 1923
Db 200 CTGCTGCAATGAGTACTTGAAGAGCAAGTGTGATCCACCGGAGACTGTGGCTGTGGA 259
Qy 1924 ACTGCTGTGAGAGAGAAATGTCTGTAAG-ATCAGTACTTT-GGAGATGTCCGAGA 1981
Db 260 ACTGCTGTGAGAGAGAAATGTCTGTAAGAAATCACTTGTGGGAGATGTCCGAGA 319

Qy 1982 GGAAGCC---GATGGGCTATGACGCTCAGGGGCTCTCAG 2020
Db 320 GGAAGCCGATTTGGGGGTCTATGACGCTCAGGGGCTCTCAG 361

RESULT 6

US-09-948-802-5
Sequence 5, Application US/09948802
Patent No. 6465232
GENERAL INFORMATION:
APPLICANT: ROBINSON, KEITH E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
FILE REFERENCE: NMT-090
CURRENT APPLICATION NUMBER: US/09/948, 802
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 09/387, 212
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentlin Ver. 2.0
SEQ ID NO 5
LENGTH: 361
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-948-802-5

Query Match 11.1%; Score 296.2; DB 3; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.1e-43;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

Qy 1684 CGAGATCTCGAAGCAGTACAGCCCAACATCGGTCTCATTTGATGTGACACC 1743
Db 20 CNAAGATCTCGAAGCAGTACAGCCCAACATCGGTCTCATTTGATGTGACACC 79
Qy 1744 AGAAGACCCCATCTATCATGTGAGAGCTTGTGAGGGGGGCACTTCTGACCTTCC 1803
Db 80 AGAAGACCCCATCTATCATGTGAGAGCTTGTGAGGGGGGCACTTCTGACCTTCC 139
Qy 1804 TCCGCAACGAGGGGGGCGGCTCGGGGTGAAGACTCTGTGCAAGATGTGGGAGATGAC 1863
Db 140 TCCGCAACGAGGGGGGCGGCTCGGGGTGAAGACTCTGTGCAAGATGTGGGAGATGAC 199
Qy 1864 CTGCTGCAATGAGTACTTGAAGAGCAAGTGTGATCCACCGGAGACTGTGGCTGTGGA 1923
Db 200 CTGCTGCAATGAGTACTTGAAGAGCAAGTGTGATCCACCGGAGACTGTGGCTGTGGA 259
Qy 1924 ACTGCTGTGAGAGAGAAATGTCTGTAAG-ATCAGTACTTT-GGAGATGTCCGAGA 1981
Db 260 ACTGCTGTGAGAGAGAAATGTCTGTAAGAAATCACTTGTGGGAGATGTCCGAGA 319
Qy 1982 GGAAGCC---GATGGGCTATGACGCTCAGGGGCTCTCAG 2020
Db 320 GGAAGCCGATTTGGGGGTCTATGACGCTCAGGGGCTCTCAG 361

RESULT 7

US-09-949-016-12110
Sequence 12110, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL0010307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241, 755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237, 768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231, 498

```

: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: / SOFTWARE: FASTSEQ for Windows Version 4.0
: / SEQ ID NO 12110
: / LENGTH: 19152
: / TYPE: DNA
: / ORGANISM: Human
: /
: JS-09-949-016-12110

```

Query Match	8.4%	Score 223.4;	DB 4;	Length 19152;
Best Local Similarity	99.6%	Pred. No. 2.4e-30;		
Matches 224;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;

Qy	60	CGAAGACGCACTATGGGGCTTCTCTCTGAGCTGTGAGAGCCCCAGGGGCGCA	CGGGGTCCTG	119
Db	18729	CAGAACGCACTATGGGCTTCTCTCTGAGCTGTGAGAGCCCCAGGGGCGCA <td>CGGGGTCCTG</td> <td>18788</td>	CGGGGTCCTG	18788
Qy	120	CAGCAATGCAAGAGGCGCAGACTTCGTCTACTGAGGGCATGAAATGTGATGGCCAG		179
Db	18789	CAGCAATGCAAGAGGCGCAGACTTCGTCTACTGAGGGCATGAAATGTGATGGCCAG		18848
Qy	180	CGGGTCAAGAGTGCACAGGAGTATGCAAGACTCTTCACCACTGTCCCTGACAGACACT		239
Db	18849	CGGGTCAAGAGTGCACAGGAGTATGCAAGACTCTTCACCACTGTCCCTGACAGACACT		18908
Qy	240	GGGGGCCAGAGCCGGGGCCATCAACCCCTGACACAGGCCCATCAGTACAG	284	
Db	18909	GGGGGCCAGAGCCGGGGCCATCAACCCCTGACACAGGCCCATCAGTACAG	18953	

RESULT 8
US-09-949-016-15795
; Sequence 15795, Application US/09949016

```

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO: 15795
LENGTH: 19153
TYPE: DNA
ORGANISM: Human
US-09-949-016-15795

```

	Query Match	8.4%	Score 22.3	DB 4	Length 19153
	Best Local Similarity	99.6%	Pred. 2.4e-30		
	Matches 224	Conservative	0	Mismatches 1	Indels 0
				Gaps 0	
QY	60	CGGAACGACATATGGGCTTCTCTTCTGAGCTGTGACGCCCGGAGCCACGGGGTCTCTG	119		
Db	18729	CAGAACGACCTAATGGGCTTCTCTTCTGAGCTGTGACGCCCGGAGCCACGGGGTCTCTG	18788		
QY	120	CAGCAATATGACGAGAGCGCAGCTTCTGTCTACTGAGAGGAGATGAAATGTGATGGCCAG	179		
Db	18789	CAGCAATATGACGAGAGCGCAGCTTCTGTCTACTGAGAGGAGATGAAATGTGATGGCCAG	18848		
QY	180	CGGCTCAAGATGACAGGAGATATGACGATCTTCAACACATGTGCTCGTAGAGACGT	239		
Db	18849	CGGCTCAAGATGACAGGAGATATGACGATCTTCAACACATGTGCTCGTAGAGACGT	18908		
QY	240	GGGGGCCAGAGCCGGGCGCATCAGCCCTTGACAGGCCCATCAGTCTAG	284		

Db 18909 GGGGGCCAGACCGGGCCATCAGCCCTGACAGGCCCATCAGTCAg 18953

RESULT 9
US-09-814-915A-82
; Sequence 82, Application US/09814915A

Query Match	7.4%	Score 198	DB 4	Length 3370
Best Local Similarity	54.7%	Pred. No. 5e-26		
Matches 421	Conservative	0	Mismatches 340	Indels 9
				Gaps 1

OY	1522	GGGTCCTGAACATGAGGACCTGGGTGTGGGTGACGAGATTGACGGGGGAACTTTGGCG	1581
Db	1967	GGGAGCTTGATCCAGCTGGGCTGATGTGTGACACTCTCATAGAGAAAGAGAGATTGGGG	2026
OY	1582	AAAGTTGACGGGAGCGCTTCGAGCC-----GACAAACACCTGTGTGGCGGTGAAGT	1632
Db	2027	AAAGTTATCGAGGGGACCCCTCAGGCTCCCGACGGACGAGACTGCAAGACTGTGGCCATTAAAG	2086
OY	1633	CTTGTGAGAGAGCGCTCCACCTGACCTCAAGGCCAAAGTTTCAACAGAAACGGAGATCC	1692
Db	2087	CCTTAAAGACACATCCCGGAGGTGGCAAGTGTGAACTTCTTTCGAGAGGCACTATCA	2146
OY	1693	TGAACAGTACAGCACCACCAATCGTGCCTCTCATTTGGTGTCTGCAGCCAGAGAGAGC	1752
Db	2147	TGGGCGAAGTTTGGCACCACCGCATATTCTGCATCTGMAAGGCGTGTCAACAAAGGAAAGC	2206
OY	1753	CCATTCAATCGTCATGAGACTTGTGACAGGGGGGCACTTCTCGACCTTCTCCGACAG	1812
Db	2207	CGATTATGATCATCAAGAAATTTATGAGAAATGACACCCCTTGATGCTTCTCTGAGGAGC	2266
OY	1813	AGGGGGCCGCGCTGGGGGTGAAAGACTGTGCTGCAGATGTTGGGGGAAATGACACTGCGCA	1872
Db	2267	GGGAGGACCAAGTGTGCTCCCTGGGCAAGTATGGCCATGCTGACAGGCGATAGCATCTGGCA	2326
OY	1873	TGAGTACCTGGAGAGCAAGTCTGCACTCCACCGGACCTGCTGCTCGAACTGCTCG	1932
Db	2327	TGAATTAACCTCAGTATATCACAATTATGTCCACCGGAGACCTGGCTGCAGAAACATCTTTGG	2386
OY	1933	TGACAGAGAAAGATGCTCGAAGATCAATGACTTTGGGATGTCCGAGAGAAAGCCGATG	1992
Db	2387	TGAATCAAAACCTGTGCTGCAAGGTGTGACTTTGGCCCTGCACTGCTCTCTCGATATGACT	2446
OY	1993	GGGTCTATGACAGCTCAGGGGGGCTCAGACAAGTCCCGTGAATGGAGCCGCACTTGAG	2052
Db	2447	TTGATGGGCACTATCAGAAACCCAGGGAGGAAAGATCCCTATCCGTTGGACAGCCCTBAAG	2506
OY	2053	CCCTTAACTACGGCCGCTACTCTCTCCGAAAGGACGTGTGAGCTTTGGCATCTTGCTCT	2112
Db	2507	CCATTGCCCATGGAATCTTCAACACAGCCAGCATGTGTGAGACTTTGGGATTTGTGATGT	2566
OY	2113	GGGAGACCTTCAAGCCGTGGGGGCTCCCGCTATCCCAACCTCAGCAATCAGACAGACGGG	2172
Db	2567	GGGAGTGTGCTGTGGGGGCAAGGCTTATGGGAGATGAGCATCAGAGGCTTATGA	2626

2173 AGTTGTGAGAAAGGGGGCCGTCGCTCCGACAGAGCTGTCTGATGCGGTGTA 2232
2627 AGAGCATTAAGATGGGTACCGGTTGCCCTCTGTGACCTGCCCCCTCTGTATG 2686
2233 GGCCTATGAGAGAGTGGGCTTATAGCTGGGAGCGGCCCACTTC 2282
2687 AGCTATGAAGAACTGCTGGGCAATATGACCTGCGCGGCCCACTTC 2736

RESULT 10
US-09-949-016-628
; Sequence 628, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 628
; LENGTH: 3921
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-628

Query Match 6.9%; Score 185.4; DB 4; Length 3921;
Best Local Similarity 55.1%; Pred. No. 8.9e-24;
Matches 385; Conservative 0; Mismatches 311; Indels 3; Gaps 1;

1617 CTGTGCGGCTGAAGTCTTGTGAGAGAGCTTCCCACTGACCTCAAGGCCAAGTTCTA 1676
2037 CCGGTGGCCATCAAGCGCTGAAGAGCCGGCTACACAGAGAGCAGAGTGCCTTC 2096
1677 CAGGAAGCAGAGATCTCTGAAGAGTACAGCAACCCCAATCGTGTCTCATTTGTGTC 1736
2097 GGGAGGCGGCACTCATGAGGCGCACTTCAAGCAACATCATCTCCCTGAGAGGCGTC 2156
1737 TGCACCCAGAGAGCCATCTCATCTGTCATGATGAGACTTGTGAGAGGAGCGACTTC 1796
2157 ATCTCAAAATCAAGGCCATGATGATCATCTGATGATGATGATGATGATGATGATGAT 2216
1797 ACCTTCTCCGACGAGAGGGGCGCGCTGCGGGTGAAGAATCTGTGTCAGATGTGGG 1856
2217 AAGTTCCTTCGGAGAGAGATGAGGCGATTCAGGCTGCTGAGCTGGTGGCATGTGCG 2276
1857 GATGAGCTGCTGAGTGAAGTACCTGAGAGAGCAATGCTGATCCACCGGAGCTGGCT 1916
2277 GGCATGCAAGCTGAGTGAAGTACCTGAGCAACATGAATATGATGATGATGATGATGAT 2336
1917 GCTCGAAGCTGCTGAGTGAAGTGAAGAAATGCTGTAAGATCAAGTATTTGGATGTCC 1976
2337 GCGCGAAGATCTCTGATCAAGCAACCTGATGATGATGATGATGATGATGATGATGATG 2396
1977 CGAAGAGAAAGCCGATGGGCTTATGAGAGCT---CAGGGGGCTTCAAGACAGTCCCGTG 2033
2397 CGGCTGCTGAGAGAGCAACCCGAGGCACTTACACACAGCTGGCGGAAAGATCCCATC 2456
2034 AAGTGAACCGACCTGAGAGGCTTAACTACGCGGCTTACCTTCCGAAAGCAAGCTGTGG 2093
2457 CGCTGAGACGCGCCGAGGCACTTCTTACCGGAAGTTCACTTCCGAGCGAGCTGTGG 2516
2094 AGCTTTGGCATTTGCTGCTGAGAGACCTTCAAGCTGGGGGCTCCCGCTATCCCAACTTC 2153
2517 AGCTTTGGCATTTGCTGCTGAGAGACCTTCAAGCTGGGGGCTCCCGCTATCCCAACTTC 2576

2154 AGCATACAGACACAGGAGTTGTGAGAGAGGGGGCGCTCTGCTCCCAAGACTG 2213
2577 TCCAAACAGAGAGTATGAAGAACCATCATATGATGCTTCCGGCTCCCAACCCATGAC 2636
2214 TGTCTGATGCGGCTGTTGAGCTTCAAGAGAGAGTGTGGGCTTATGAGCTGGGAGCGG 2273
2637 TGCCCTCGGCATCTTACAGAGTTCATGATGATGATGATGATGATGATGATGATGATG 2696
2274 CCAGCTTCAGACCATCTTACAGAGAGCTGCAAGACTC 2312
2697 CCAGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2735

RESULT 11
US-09-949-016-2120
; Sequence 2120, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2120
; LENGTH: 3921
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2120

Query Match 6.9%; Score 185.4; DB 4; Length 3921;
Best Local Similarity 55.1%; Pred. No. 8.9e-24;
Matches 385; Conservative 0; Mismatches 311; Indels 3; Gaps 1;

1617 CTGTGCGGCTGAAGTCTTGTGAGAGAGCTTCCCACTGACCTCAAGGCCAAGTTCTA 1676
2037 CCGGTGGCCATCAAGCGCTGAAGAGCCGGCTACACAGAGAGCAGAGTGCCTTC 2096
1677 CAGGAAGCAGAGATCTCTGAAGAGTACAGCAACCCCAATCGTGTCTCATTTGTGTC 1736
2097 GGGAGGCGGCACTCATGAGGCGCACTTCAAGCAACATCATCTCCCTGAGAGGCGTC 2156
1737 TGCACCCAGAGAGCCATCTCATCTGTCATGATGAGACTTGTGAGAGGAGCGACTTC 1796
2157 ATCTCAAAATCAAGGCCATGATGATCATCTGATGATGATGATGATGATGATGATGATG 2216
1797 ACCTTCTCCGACGAGAGGGGCGCGCTGCGGGTGAAGAATCTGTGTCAGATGTGGG 1856
2217 AAGTTCCTTCGGAGAGAGATGAGGCGATTCAGGCTGCTGAGCTGGTGGCATGTGCG 2276
1857 GATGAGCTGCTGAGTGAAGTACCTGAGAGAGCAATGCTGATCCACCGGAGCTGGCT 1916
2277 GGCATGCAAGCTGAGTGAAGTACCTGAGCAACATGAATATGATGATGATGATGATGATG 2336
1917 GCTCGAAGCTGCTGAGTGAAGTGAAGAAATGCTGTAAGATCAAGTATTTGGATGTCC 1976
2337 GCGCGAAGATCTCTGATCAAGCAACCTGATGATGATGATGATGATGATGATGATGATG 2396
1977 CGAAGAGAAAGCCGATGGGCTTATGAGAGCT---CAGGGGGCTTCAAGACAGTCCCGTG 2033
2397 CGGCTGCTGAGAGAGCAACCCGAGGCACTTACACACAGCTGGCGGAAAGATCCCATC 2456
2034 AAGTGAACCGACCTGAGAGGCTTAACTACGCGGCTTACCTTCCGAAAGCAAGCTGTGG 2093

DB 2457 CGCTGACCGCCCGGAGGACCATTTCTCAACCGAATTCACTCTGCAGCAGCTGTGG 2516
QY 2094 AGCTTGGGATCTGTCTCTGGGAGACCTTACGCTTGGGGGCTCCCTTATCCCACTTC 2153
DB 2517 AGCTTGGGATCTGTCTCTGGGAGGAGTGAACCTTATGCGAGCGCCCTTACGGAGTTG 2576
QY 2154 AGCAATCAGACAGACGAGGAGTTTGTGAGAAAGGGGGCCGTCTGCTCCGCGCAAGCTG 2213
DB 2577 TCACACACGAGGTATGAAAGCATCAATGATGCTTCCGGCTCCCAACCCATGAGAC 2636
QY 2214 TGTCTGATGCGGTGTTCAGGCTCATGAGCAGTGTGGGCTTATGAGGCTTGGGAGCGG 2273
DB 2637 TGCCCTCGGACCTACCAAGCTCATGATGAGTGTCTGGCAGCAGAGCGGCGCGC 2696
QY 2274 CCCAGCTTACGACCATCTTACCAAGAGCTGCAAGATC 2312
DB 2697 CCCAGTGTGCTGACATGCTGACAGATCTGGAACAAGCTC 2735

RESULT 12
US-09-949-016-5662
Sequence 5662, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5662
LENGTH: 3386
TYPE: DNA
ORGANISM: Human
US-09-949-016-5662

Query Match
Best Local Similarity 54.7%; Pred. No. 7.1e-22;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;
Score 174.6; DB 4; Length 3386;

QY 1569 GGGAACTTTGGCGAAGTGTTCAGCGGAGCGCTGCGAGCGCAACAACCTGTGGCGGTG 1628
DB 1386 GGGGAGGTCTATGAGGTGTCTACACAATCAACAAGGGAGAAATCAATGAGCTGTC 1445
QY 1629 AAGCTTGTGAGAGAGCGCTCCCACTGACCTCAAGGCCAAGTTTCTACAGAAAGGAG 1688
DB 1446 AAGACCTGCAAGAAAGACTGCACTGAGCAACAAGAGAAAGTTCAATGAGCAGGAGTG 1505
QY 1689 ATCTTGAAGAGAGACGACCGCAACCTGCTGCTGCTATGTGTCTGCAACCCGAGAG 1748
DB 1506 ATATGAAAGAACTGACCAACCGCACATCTGTAAGCTGATGCGCAAT---CATTTAAGAG 1562
QY 1749 CAGCCCATCTACATCTCATGAGAGCTTGTGCAAGGGGGCGAATCTTCTGACCTTCTCCG 1808
DB 1563 GAGCCCACTGATCATCATGAAATGTATCCCTATGAGGAGGTGGGCCCATCTACCTGGAG 1622
QY 1809 ACGAGAGGGGCGCGCTGCGGTGAAGACTCTGCTGCAAGTGTGAGGAGATGCAAGCTGCT 1868
DB 1623 CGGAACAAGAACTCCCTGAAGGTGCTCAACCTGCTGTAATCACTGACAGATATCAAA 1682
QY 1869 GCGATGAGAGTCTGAGAGGAGAGTGTGATCCACCGGAGACTGTGCTGGAACTGC 1928
DB 1683 GCGATGCGCTTACCTGAGAGCATCACTGCTGCAACGGGACATTCGTCTCCGAAACATC 1742
QY 1929 CTGCTGACAGAGAAATGTCTGTAAGATCAGTGAATTTGGAGTGTCCCGAGAGAGCC 1988

DB 1743 CTGTGTGCTCTCCCTGAGGTGTGAAGCTGGGGGACTTTGTCTTCCCGGTACATTGAG 1802
QY 1989 GATGGGCTTATGACAGCTCAGGGGGCTTCAGACAAGTCTCCGTGAAGTGAACCGACCT 2048
DB 1803 GAGGAGGACTAT---TACAAAGCCCTGTGTGACTGTCTTCCCATCAAAAGATGTCCCA 1859
QY 2049 GAGGCCCTTAACCTAGCGCGCTACTCTCCGAAAGCAGTGTGAGCTTTGGCATCTTG 2108
DB 1860 GAGTCCATTAACTTCCGAGCTTCAAGACGACAGCATGAGTGTGATGTTCCCGGTGTC 1919
QY 2109 CTCTGGAGACCTTCAAGCTGCGGGGCTCCCTTATCCCAACTGACAAATCAGACAGACA 2168
DB 1920 ATGTGGAGATCTGAGCTTGTGGAAGCAGCCCTTCTTGTGCTGGAGAAACAAGATGTC 1979
QY 2169 CGGAGTTTGTGAGAAAGGGGGCCGTCTGCTCCCTGCGCAGACTGTGTCTGATGCGCTG 2228
DB 1980 ATCGGGGTGTGGAAGAAAGAGACCGGCTGCCAAGCTGATCTGTTCACCGGTCTT 2039
QY 2229 TTCAGGCTCATGAGCAGTGTGGGCTTATGAGCTTGGGACGCGCCGACCTTCAAGC 2285
DB 2040 TTATCCCTCATGACCGCTGCTGTGGACTAGACACCCCATGATGACCGGCGGCTTCAAC 2096

RESULT 13
US-08-357-642A-2
Sequence 2, Application US/08357642A
Patent No. 5837524
GENERAL INFORMATION:
APPLICANT: Sima Lev
APPLICANT: Joseph Schlessinger
TITLE OF INVENTION: PYK2 RELATED PRODUCTS
TITLE OF INVENTION: AND METHODS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Suite 4700
STREET: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,642A
FILING DATE: December 15, 1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
ATTORNEY/AGENT INFORMATION:
FILING DATE:
NAME: Marburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic
US-08-357-642A-2

Query Match
Score 174.6; DB 2; Length 3416;

Best Local Similarity 54.7%; Pred. No. 7.1e-22;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

1569 GGGAACTTTGGGAGAGTGTTCAGCGGACGCTGCGAGCCGACAAACACCTGTGGCGGTG 1628
1416 GGGAGAGTCTATGAGAGGTGTCTACACAATACAAAGGGAGAAATCAATGATGTGTC 1475
1629 AAGCTTTTCGAGAGACGCTCCACCTGACCTTAAGGCCAAGTTTCTACAGAGGAGAG 1688
1476 AAGACCTGGAAGAAAGACTGCACTCTGAGCAACAGAGAAAGTTCAATGAGCGGAGAGT 1535
1689 ATCTGAGAGAGTACAGCAACCCCAACATGTCGCTCACTGATGTCGTCGACCCAGAG 1748
1536 ATCATAGAGAACTTCAACACCCGACATCTGTAAGCTGATCCGCACT--CATTTAAGAG 1592
1749 CAGCCCATCTACATCTGATGAGCTTGTGTCAGAGGGGCGCACTTCTGACCTTCTCCG 1808
1593 GAGCCACCTGATCATCATGAAATTGTATCCCTATGGGAGCTGGGCCACTAGAG 1652
1809 ACGGAGGGGCGCGCTGCGGCTGAAAGACTTGTCTGCAAGTGTGGGGAGTCACTGCT 1868
1653 CGGAAACAGAACTCCCTGAAAGGTCTCACCTCTGTCTGATCACTGACAGATATGCAAA 1712
1869 GGCATGAGTACCTGAGAGCAAGTGTGATCCACCGGACCTGTGCTGCGGAATG 1928
1713 GGCATGCGCTTACCTGAGAGCACTCACTGCTGACAGAGGACATGCTGCTCCGAAACATC 1772
1929 CTGTGACAGAGAAAGATGTCTGAAGATCACTGATGATGTCGATGTCGCGAGAGAGCC 1988
1773 CTGTGAGCTCCCTGATGATGTGAGAGCTGGGGAATTTGTGTTCCGGTACATTTAG 1832
1989 GATGGGCTCTATGACCTCTCAGGGGCTCTCAACAAAGTCCCGTGAAGTGAACCGACCT 2048
1833 GACGAGAGACTAT--TACAAAGCCTCTGTGATCTGCTCCCATCAAAAGGATGTCCCA 1889
2049 GAGGCGCTTAACTACGGCGCTACTCTCCGAAAGGAGCGTGTGAGCTTTGGCATCTTG 2108
1890 GAGTCCATTAACTTCGAGCGCTTCAAGACAGCAAGTGAAGTGTGAGTGTGCGGATGTC 1949
2109 CTCTGAGAGACCTTTCAGCGCTGAGGGGCTCCCTATCCCACTCAGCAATCAGAGACA 2168
1950 ATGTGGAGATCTCTGAGCTTTGGGAGAGCAAGCTTCTTCTGTGGTGAAGAAAGATGTC 2009
2169 CCGGAGTTTGTGAGAAAGGGGCGCTCTGCTGCTGCCAGAGCTGTGTCTGATGCTG 2228
2010 ATCGGGGTGTGAGAAAGGAGACCGGCTGCCAAGCTGATCTGTCCACCGGTCTT 2069
2229 TTCAGGCTCATGAGAGTGTGGGCTTATGAGCTGGGAGGAGGCGCCAGCTTCAG 2285
2070 TATACCTCATGACCGCTGTGTGGACTACGACCCAGTGAACGCGCCCGCTTACCC 2126

RESULT 14
US-08-460-626-2

Sequence 2, Application US/08460626
Patent No. 5837815
GENERAL INFORMATION:
APPLICANT: SIMA LEV
APPLICANT: JOSEPH SCHLESSINGER
TITLE OF INVENTION: PK-2 RELATED PRODUCTS AND
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: Storage

COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,626
FILING DATE: June 2, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/357,642
FILING DATE: December 15, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 211/121
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-626-2

Query Match 6.5%; Score 174.6; DB 2; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.1e-22;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

1569 GGGAACTTTGGGAGAGTGTTCAGCGGACGCTGCGAGCCGACAAACACCTGTGGCGGTG 1628
1416 GGGAGAGTCTATGAGAGGTGTCTACACAATACAAAGGGAGAAATCAATGATGTGTC 1475
1629 AAGCTTTTCGAGAGACGCTCCACCTGACCTCAAGGCCAAGTTTCTACAGAAAGGAG 1688
1476 AAGACCTGGAAGAAAGACTGCACTCTGAGCAACAGAGAAAGTTCAATGAGCGAGAG 1535
1689 ATCTGAGACAGTACAGCAACCCCAACATCTGTGCTGATGTCGATGTCGCGAGAG 1748
1536 ATCATAGAGAACTTCAACACCCGACATCTGTAAGCTGATGGCACT--CATTTAAGAG 1592
1749 CAGCCCATCTACATCTGATGAGCTTGTGTCAGAGGGGCGGACTTCTGACCTTCTCCG 1808
1593 GAGCCCACTGTGATCAATGAAATTGTATCCCTATGAGGAGGCTGGCCACTAGAG 1652
1809 ACGGAGGGGCGCGCTGAGGGGTAAGACTGTGCTGACATGATGATGGGAGTGCAGCTG 1868
1653 CGGAAACAGAACTCCCTGAAAGGTCTCACCTCTGTCTGATCACTGACAGATATGCCAA 1712
1869 GGCATGAGTACCTGAGAGCAAGTGTGATCCACCGGACCTGTGCTGCGGAATGTC 1928
1713 GGCATGCGCTTACCTGAGAGCACTCACTGCTGTCAGAGGAGCACTGCTGCGGAACATC 1772
1929 CTGTGACAGAGAAAGATGTCTGAAAGATCACTGATGATGTCGATGTCGCGAGAGAGCC 1988
1773 CTGTGAGCTCCCTGATGATGTGTAAGCTGGGAGCTTTGTGTTCCGGTACATTTAG 1832
1989 GATGGGCTATGAGAGCTCAGGGGCTCAAGCAAGTCCCGGTGAAGTGAACCGACCT 2048
1833 GAGGAGACTAT--TACAAAGCTTGTGATCTGCTCTCCCATCAAAAGATGTGCCA 1889
2049 GAGGCGCTTAACTACGCGCTACTCTCCGAAAGGAGCGTGTGAGCTTTGGCATCTTG 2108
1890 GAGTCCATTAACTTCGAGCGCTTCAAGACAGCAAGTGAAGTGTGAGTGTGCGGATGTC 1949
2109 CTCTGAGAGACCTTTCAGCTTGGGCGCTCCCTATCCCACTCAGCAATCAGAGACA 2168
1950 ATGTGGAGATCTGTGCTTTGGGAGAGCAAGCTTCTTCTGTGCTGAGAAACAGAGATGTC 2009
2169 CCGGAGTTTGTGAGAAAGGGGCGCTGCTGCTGCCAGAGAGCTGTGCTGATGCTG 2228
2010 ATCGGGGTGTGAGAAAGGAGACCGGCTGCCAAGCTGATCTGTCTCACCGGTCTT 2069

QY 2229 TTCAGGCTCATGAGAGAGTGTGGGCTTATGAGCCTTGGGAGAGCCGACGTTACAG 2285
DB 2070 TATACCTCATGACCCGCTGTGGAGCTAGACCCCACTGACCGGCGCCGCTTACCC 2126

RESULT 15
US-09-016-434-1483
Sequence 1483, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1483:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g988304
US-09-016-434-1483

Query Match 6.5%; Score 174.6; DB 4; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.1e-22;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

QY 1569 GGGAACTTGGCGAAGTGTTCAGCGGAGCGCTTCCGAGCCGCAACAACCTGTGGCGGTG 1628
DB 1416 GGGGAGGTCTATGAGAGGTGTCTACACAAATCACAAAGGGGAGAAATCAATGTAGCTGTC 1475
QY 1629 AAGCTTGTTCGAGAGACGCTCCCACTGACCTCAAGGCCAAGTTTCTACAGGAAGGAGG 1688
DB 1476 AAGACCTGCAAGAAAGCTGCACTTGGACCAAGAGAGAGTTCAATGAGCGGAGGAGTG 1535
QY 1689 ATCTGAAGAGTACAGCCACCCCAATCGTGCCTCTCATTTGATGTGCAACCAAG 1748
DB 1536 ATCATGAGAGAACTCTGACCAACCGGCAATGTGAGTGAATGGCAT---CATGGAAG 1592
QY 1749 GAGCCCATCTACATCGTCATGAGAGCTTGTGCAAGGGGGCGACTTCTGACCTTCTCCG 1808
DB 1593 GAGCCCACTGATTCATGAGAAATGTATCCTATGGGAGGCTGGGCGCACTACCTGAG 1652

QY 1809 ACGAGGGGGCCCGGCTGCGGGTGAAGACTTGTGTCAGAGATGATGGGAGATGCACTGTG 1868
DB 1653 CGGAACAAGAACTCCCTGTAAGGTGCTACACCCCTGTGTGTACTCACTGAGATATGCAAA 1712
QY 1869 GGCATGAGATACCTGGAAGAGCAAGTGTGATTCACCGGAGACTGTGCTGGAACTGC 1928
DB 1713 GCCATGGCCCTACTGGAAGAGCTCAACTGCGTGCACAGGGAATTTGCTCCGGAACATC 1772
QY 1929 CTGTGACAGAGAAATCTCTGAAGATCAAGTGACTTTGGGATGTCCGAGAGGAAGCC 1988
DB 1773 CTGTGGCCCTCCCTGAGTGTGGAAGCTGGGGGACTTGTGCTTCCGGTACATTGAG 1832
QY 1989 GATGGGCTATGACAGCTCAGAGGGGCTCAGACAAGTCCCGTGAAGTGAACCCGACCT 2048
DB 1833 GACGAGACTAT---TACAAAGCTCTGTGACTGTCTCCCATCAAAATGATGTCCCA 1889
QY 2049 GAGGCCCTTAACCTACGCGGCTACTCTCCGAAAGCGAGTGTGAGCTTTGGCATCTTG 2108
DB 1890 GAGTCATTTAACTTCCGAGGCTTCAAGACGACAGTGAAGTGTGATGTTCCCGGTGTC 1949
QY 2109 CTGTGGAGACCTTCAAGCTGGGGGCTCCCTTATCCCAACTCAGCAATGAGCAGACA 2168
DB 1950 ATGTGGAGATCTGAGCTTTGGGAAAGCAGCCCTTCTGTGCTGAGAAAGATGTTC 2009
QY 2169 CGGAGTTTGTGAGAGAGGGGGCGGTCTGCGCTGCGCCAGAGCTGTGCTGATGCCGTG 2228
DB 2010 ATCGGGGTCTGAGAGAGAGAGAACCGGCTGCCAAGGCTGATCTGTGTCAACGGATCTT 2069
QY 2229 TTCAGGCTCATGAGAGAGTGTGGGCTATGAGCCTTGGGAGAGGCGCCAGCTTACAG 2285
DB 2070 TATACCTCATGACCCGCTGTGGAGCTAGACCCCACTGACCGGCGCCGCTTACCC 2126

Search completed: March 19, 2005, 14:24:13
Job time : 442.254 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 03:22:22 ; Search time 7723.14 Seconds
(without alignment)
13179.075 Million cell updates/sec

Title: US-10-660-763-1

Perfect score: 2674

Sequence: 1 tccgggggtccgaccgggccc.....aaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hlc:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1806.8	67.6	2461	3	CR624741 full-length
2	1189	44.5	1189	3	CR593957 full-length
3	923.6	34.5	936	5	BX356089 BX356089
4	908.8	34.0	980	5	BX356088 BX356088
5	897.6	33.6	3373	3	AK004587 Mus muscu
6	885.4	33.1	936	1	AL542976 AL542976
7	875.8	32.8	1029	4	BM456755 AGENCOURT
8	863	32.3	863	3	CR622908 full-length
9	855.2	32.0	989	5	BX359009 BX359009
10	836.6	31.3	861	1	AL540776 AL540776
11	821.6	30.7	968	5	BX359010 BX359010
12	805	30.1	848	1	AL540775 AL540775
13	804.4	30.1	939	5	BX381726 BX381726
14	791.6	29.6	913	1	AL569718 AL569718
15	771.6	28.9	948	5	BQ708270 AGENCOURT
16	731	27.3	732	5	BQ015500 UI-H-E11-
17	716.8	26.8	922	5	BUI56855 AGENCOURT
18	710.6	26.6	1114	4	BM543870 AGENCOURT
19	703.2	26.3	847	5	BUI84736 AGENCOURT
20	681.2	25.5	688	5	BQ016080 UI-H-DT1-
21	679	25.4	734	1	AI862686 wj27903.x
22	671.6	25.1	735	5	BQ711139 AGENCOURT
23	669.4	25.0	973	5	BX327963 BX327963
24	652.8	24.4	802	5	BP157303 BP157303

25	640.2	23.9	787	6	CB990282 AGENCOURT
26	640	23.9	659	5	BM929822 UI-E-E11-
27	629	23.5	629	7	CV030188 full
28	628.8	23.5	1077	5	BM922603 AGENCOURT
29	613.8	23.0	882	6	BY704806 BY704806
30	613.4	22.9	880	5	BX381725 BX381725
31	606	22.7	638	5	BQ807439 NISC JK04
32	603	22.5	629	4	BM684264 UI-E-E11-
33	600.8	22.5	652	7	CN298839 170004240
34	600	22.4	602	6	CA449888 UI-H-E11-
35	599.4	22.4	962	5	BQ679512 AGENCOURT
36	597	22.3	938	5	BQ712558 AGENCOURT
37	593.8	22.2	678	2	AW130256 xE47B03.x
38	588	22.0	782	7	CK357411 AGENCOURT
39	587.6	22.0	911	5	BUI91031 AGENCOURT
40	581.6	21.8	812	7	CK788677 AGENCOURT
41	578.4	21.6	580	5	BP332887 BP332887
42	576.4	21.6	674	7	CK943999 4068060 B
43	574.8	21.5	912	5	BQ924403 AGENCOURT
44	568	21.2	582	5	BP256574 BP256574
45	551.2	20.6	662	2	AW701492 uq86h02.y

ALIGNMENTS

RESULT 1	CR624741	2461 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	CR624741				
DEFINITION	full-length cDNA clone CS0DI069YK21 of Placenta Cot 25-normalized				
ACCESSION	CR624741.1	GI:50505548			
VERSION	CR624741.1	GI:50505548			
KEYWORDS	HTC; CNSLT; cDNA				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifeitech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paraday Avenue 2 (bases 1 to 2461)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :				
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..2461				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CS0DI069YK21"				
	/issue_type="Placenta Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
Query Match	67.6% ; Score 1806.8 ; DB 3 ; Length 2461 ;				
Best Local Similarity	85.3% ; Pred. No. 2.1e-31 ;				
Matches 2205 ; Conservative	0 ; Mismatches 2 ; Indels 377 ; Gaps 2 ;				
Oy	41 CCAGAGCAGCTGCCGTGCGGAGCAGCAGCTTGGCTTCTTCTGAGCTGTGAGCC 100				
Db	1 CCAGAGCAGCTGCCGTGCGGAGCAGCAGCTTGGCTTCTTCTGAGCTGTGAGCC 60				

101 CCAGGCGCAGGGGCTCTGAGCAAAATGAGAGGCCAGCTTCTCTACTGAGGGCAAT 160
102 CCAGGCGCAGGGGCTCTGAGCAAAATGAGAGGCCAGCTTCTCTACTGAGGGCAAT 120
161 GAGAAATGAGATGAGCCAGCGAGTCAAGATGACAGGAGATGACAGACTGCTTACCA 220
121 GAGAAATGAGATGAGCCAGCGAGTCAAGATGACAGGAGATGACAGACTGCTTACCA 180
221 CATGTCCTGAGAGACAGTGGGGGCGAGAGCCCGGCTATCAGCCCTGAGACCCCATCAG 280
181 CATGTCCTGAGAGACAGTGGGGGCGAGAGCCCGGCTATCAGCCCTGAGACCCCATCAG 240
281 TCAGTCTGAGAGTCAAGATGAGAGCCAGCAAACTGAGGGCTGAGAGCCGCTGAGAGAG 340
241 TCAGTCTGAGAGTCAAGATGAGAGCCAGCAAACTGAGGGCTGAGAGCCGCTGAGAGAG 300
341 CGCAGAGATCTGAACTCAGGGGCCCTGAGCAAGCTGAGCTGCTATCCGGGAAACGGA 400
301 CGCAGAGATCTGAACTCAGGGGCCCTGAGCAAGCTGAGCTGCTATCCGGGAAACGGA 360
401 GCAGCTTCCGAGACCTTACAGCGAGAGTGGGAGAGCTGAGAGAGAGCTTACCAAGAC 460
361 GCAGCTTCCGAGACCTTACAGCGAGAGTGGGAGAGCTGAGAGAGAGCTTACCAAGAC 420
461 CCACAGCGAGACATGAGAGAGCTGAGAGCCAGTACCGAGCTTGGCAACGGGACAGTGC 520
421 CCACAGCGAGACATGAGAGAGCTGAGAGCCAGTACCGAGCTTGGCAACGGGACAGTGC 480
521 CCAGAGCAAGCGAGATCAGAGAGAGCCAGCAAAAGCAAGAGCCGTCAGACAGGCAAGGA 580
481 CCAGAGCAAGCGAGATCAGAGAGAGCCAGCAAAAGCAAGAGCCGTCAGACAGGCTAAGA 540
581 CAAGTATGAGCGAGAGCTGAGAGAGCTTCTTGTCTACCAACCGCTATGTCGTGGGCGT 640
541 CAAGTATGAGCGAGAGCTGAGAGAGCTTCTTGTCTACCAACCGCTATGTCGTGGGCGT 600
641 GCGGGCTGCGAGACTACACCAAGCAGCAGCAGCAGCTGCTGCGCGGCTGCTGCG 700
601 GCGGGCTGCGAGACTACACCAAGCAGCAGCAGCAGCTGCTGCGCGGCTGCTGCG 660
701 GTCACTGAGAGACTGAGAGAGAGTGGCTGATCTGAGAGAGATCTGAGAGAGAT 760
661 GTCACTGAGAGACTGAGAGAGAGTGGCTGATCTGAGAGAGATCTGAGAGAGAT 720
761 CCGAGAGATGAGAGAGCTGAGAGAGTGGCTGATCTGAGAGAGATCTGAGAGAGAT 820
721 CCGAGAGATGAGAGAGCTGAGAGAGTGGCTGATCTGAGAGAGATCTGAGAGAGAT 780
821 AGCTGCTGCGAGACTGAGAGAGCTGAGAGAGTGGCTGATCTGAGAGAGATCTGAGAGAG 880
781 AGCTGCTGCGAGACTGAGAGAGCTGAGAGAGTGGCTGATCTGAGAGAGATCTGAGAGAG 840
881 CGCAGCTGAGAGTCCAGCCCTGTCAGCTTTCAGTGAATGATCACTGCTTGAAGAGGTAAC 940
841 CGCAGCTGAGAGTCCAGCCCTGTCAGCTTTCAGTGAATGATCACTGCTTGAAGAGGTAAC 900
941 GCTGAGAGCTGAGAGAGTCCAGCTGAGAGAGTGGCTGATCTGAGAGAGATCTGAGAGAG 1000
901 GCTGAGAGCTGAGAGAGTCCAGCTGAGAGAGTGGCTGATCTGAGAGAGATCTGAGAGAG 960
1001 GACCTGAGAGAGTGGCTGAGAGAGTGGCTGAGAGAGTGGCTGAGAGAGTGGCTGAGAGAG 1060
961 GACCTGAGAGAGTGGCTGAGAGAGTGGCTGAGAGAGTGGCTGAGAGAGTGGCTGAGAGAG 1020
1061 GATTAGCGAGACTGAGAGAGTCCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1120
1021 GATTAGCGAGACTGAGAGAGTCCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
1121 GATTAGCGAGACTGAGAGAGTCCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1180
1081 GATTAGCGAGACTGAGAGAGTCCAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140

1181 GCTGAGAGCGAG 1240
1141 GCTGAGAGCGAG 1200
1241 CTTGGGCGCCGAG 1300
1201 CTTGGGCGCCGAG 1260
1301 GTCCTGAG 1360
1261 GTCCTGAG 1320
1361 CATCTCAGAGAGATCTTCCGCGCCAGTCTC----- 1390
1321 CATCTCAGAGAGATCTTCCGCGCCAGTCTC----- 1380
1391 ----- 1390
1381 GGTGAGAGAGCCCTGAGATGAGAGAGCTGTGTATCAACGAGGAGCATCCGAGGAGAGAG 1440
1391 ----- 1390
1441 GGTGAGAGAGCCCTGAGATGAGAGAGCTGTGTATCAACGAGGAGCATCCGAGGAGAGAG 1500
1391 ----- 1390
1501 GTCAGGCTGTCGAGT 1560
1391 -----GAACTGTATCCGACTGT 1406
1561 CCCCTCAGAGATGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1620
1407 GAAAGGAGAGAGCTTCTTCTGAGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1466
1621 GAAAGGAGAGAGCTTCTTCTGAGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
1467 CCCCTCAGAGAGAGAGT 1526
1681 CCCCTCAGAGAGAGAGT 1740
1527 CTGAACATGAGAGAGCTGT 1586
1741 CTGAACATGAGAGAGCTGT 1800
1587 TTCAGGAGAGAGCTGT 1646
1801 TTCAGGAGAGAGCTGT 1860
1647 CTCCTCAGAGAGAGAGCTGT 1706
1861 CTCCTCAGAGAGAGAGCTGT 1920
1707 CACCCCAATGAGAGAGCTGT 1766
1921 CACCCCAATGAGAGAGCTGT 1980
1767 ATGAGAGCTGT 1826
1981 ATGAGAGCTGT 2040
1827 CGGAGAGAGAGCTGT 1886
2041 CGGAGAGAGAGCTGT 2100
1887 AGCAAGT 1946
2101 AGCAAGT 2160
1947 GTCCTGAGAGAGAGCTGT 2006
2161 GTCCTGAGAGAGAGCTGT 2220
2007 TCAGGAGAGAGCTGT 2066

Db 2221 TCAAGGAGGCTCAGACAAATCCCCGTAAGATGAGACCGCACTGAGCCCTTAATCTAC--- 2277
Qy 2067 CGCTACTCTCCGAAAGCAGACGTGTGAGCTTTGACATCTTGTCTGGAGACCTTTCAGC 2126
Db 2278 ----- 2277
Qy 2127 CTGGGGGCTCCCCCTATCCCAACTCAGCAATCAGACAGACGGAGTTTGTGAGAAAG 2186
Db 2278 ----- 2277
Qy 2187 GGGGGCGCTGTGCGCCCTGACCAGAGCTGTGCTGATGACCGTGTTCAGGCTCATGAGACAG 2246
Db 2278 GGGGGCGCTGTGCGCCCTGACCAGAGCTGTGCTGATGACCGTGTTCAGGCTCATGAGACAG 2337
Qy 2247 TGCTGGGCTATGAGCTGTGGGACCGGCGCCAGCTTTCAGCACTTACAGAGAGTGCAG 2306
Db 2338 TGCTGGGCTATGAGCTGTGGGACCGGCGCCAGCTTTCAGCACTTACAGAGAGTGCAG 2397
Qy 2307 AGCATCCGAAAGCGGCAATGGTGAAGCTGGGACCCCTTCTCAAGCTGGTGGCTTTCGA 2366
Db 2398 AACATCCGAAAGCGGCAATGGTGAAGCTGGGACCCCTTCTCAAGCTGGTGGCTTTCGA 2457
Qy 2367 GGCC 2370
Db 2458 GGCC 2461

RESULT 2
CR593957 1189 bp mRNA linear HMC 21-JUL-2004
LOCUS full-length cDNA clone CS0D1008Y607 of Placenta Cot 25-normalized
DEFINITION
ACCESSION CR593957
VERSION CR593957.1 GI:50474764
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1189)
AUTHORS Li W.B., Gruber C., Jessup J. and Polyes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1189)
REFERENCE Direct Submission
AUTHORS Genoscope.
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source 1. 1189
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1008Y607"
/cruise_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 44.5%; Score 1189; DB 3; Length 1189;
Best Local Similarity 100.0%; Pred. No. 2.1e-201;
Matches 1189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1326 GGAAGACACCCACGCTGAGATCTTAAGAGCAGCATCTCAGAAATCTTCCGCCCAAG 1385

Db 1 GGAAGACACCCACGCTGAGATCTTAAGAGCAGCATCTCAGAAATCTTCCGCCCAAG 60
Qy 1386 TTCTGAACTGTATCCGATCGTGAAGGGAAGCTTCTTGAATCTCTTGTCTCATGAC 1445
Db 61 TTCTGAACTGTATCCGATCGTGAAGGGAAGCTTCTTGAATCTCTTGTCTCATGAC 120
Qy 1446 CACCTTATGAGACCCAGAGCCCTCAGCAAGAAAGTGTGTCTCTGCAAGAGGCT 1505
Db 121 CACCTTATGAGACCCAGAGCCCTCAGCAAGAAAGTGTGTCTCTGCAAGAGGCT 180
Qy 1506 GTGCCAAGACAAAGTGGTGTCTGAACTATGAGACCTGTGTGGTGAAGATTTGA 1565
Db 181 GTGCCAAGACAAAGTGGTGTCTGAACTATGAGACCTGTGTGGTGAAGATTTGA 240
Qy 1566 GGGGGAACTTGGGGAAGTGTTCAGCGGAGCGCTGCGAGCCGACAAACCTGTGTGGG 1625
Db 241 GGGGGAACTTGGGGAAGTGTTCAGCGGAGCGCTGCGAGCCGACAAACCTGTGTGGG 300
Qy 1626 GTGAAGTCTTGTGAGAGACGCTCCCACTGACCTTCAGGCCAAGTTTTCAGAGAAAGG 1685
Db 301 GTGAAGTCTTGTGAGAGACGCTCCCACTGACCTTCAGGCCAAGTTTTCAGAGAAAGG 360
Qy 1686 AAGATCTGAAGAGATGACGACCAACCCCAATCGTGCCTCATTTGTGTCTGACCCAG 1745
Db 361 AAGATCTGAAGAGATGACGACCAACCCCAATCGTGCCTCATTTGTGTCTGACCCAG 420
Qy 1746 AAGCAGCCCATCTACATCTGTATGAGACTGTGTGAGAGGGGGGAGCTTCTGACCTTCTC 1805
Db 421 AAGCAGCCCATCTACATCTGTATGAGACTGTGTGAGAGGGGGGAGCTTCTGACCTTCTC 480
Qy 1806 CGCAGGAGGGGGCGCCGCTGGGGGTGAAGACTTCTCTGCAATGTGTGGGGATGGAAGCT 1865
Db 481 CGCAGGAGGGGGCGCCGCTGGGGGTGAAGACTTCTCTGCAATGTGTGGGGATGGAAGCT 540
Qy 1866 GCTGGCATGAGATCTGTGAGAGCAAGCTGTGATCCACCGGAGCTGTGCTGCGAAG 1925
Db 541 GCTGGCATGAGATCTGTGAGAGCAAGCTGTGATCCACCGGAGCTGTGCTGCGAAG 600
Qy 1926 TGCTGTGTGACAGAGAAAGATGCTCTGAAGATCAGTGAATTTGGAGTGTCCGAGAGAA 1985
Db 601 TGCTGTGTGACAGAGAAAGATGCTCTGAAGATCAGTGAATTTGGAGTGTCCGAGAGAA 660
Qy 1986 GCCGATGGGATCTATGACAGCTTCAAGGGGGCTTCAGCAAGTCCCGTGAAGTGAACCGCA 2045
Db 661 GCCGATGGGATCTATGACAGCTTCAAGGGGGCTTCAGCAAGTCCCGTGAAGTGAACCGCA 720
Qy 2046 CCGTAGGCGCTTAACCTAGAGGCGCTACTCTCGAAGAGGAGCGTGTGAGCTTGGCATC 2105
Db 721 CCGTAGGCGCTTAACCTAGAGGCGCTACTCTCGAAGAGGAGCGTGTGAGCTTGGCATC 780
Qy 2106 TTGCTGTGAGAGACTTTCAGCTGTGGGGCTCCCTCATCCCACTCAGCAATCAGCAG 2165
Db 781 TTGCTGTGAGAGACTTTCAGCTGTGGGGCTCCCTCATCCCACTCAGCAATCAGCAG 840
Qy 2166 ACAAGGAGTTTGTGAGAAAGGGGGCGCTGTGCTTCCGCAAGTGTGTCTGATGCC 2225
Db 841 ACAAGGAGTTTGTGAGAAAGGGGGCGCTGTGCTTCCGCAAGTGTGTCTGATGCC 900
Qy 2226 GTGTTGAGCTCATGAGAGAGTGTGGGCTATGAGCTGTGGGAGGGGCCAGCTTCAGC 2285
Db 901 GTGTTGAGCTCATGAGAGAGTGTGGGCTATGAGCTGTGGGAGGGGCCAGCTTCAGC 960
Qy 2286 ACCATCTACAGAGAGCTGAGAGCATCCGAAAGCGGAGCATCGTGAAGCTGGAGCCCTT 2345
Db 961 ACCATCTACAGAGAGCTGAGAGCATCCGAAAGCGGAGCATCGTGAAGCTGGAGCCCTT 1020
Qy 2346 CTCAGCTGTGCTGTGCAAGGCTTGAAGTGTCTGACAGTGTCTCAAGGCTCATATGC 2405
Db 1021 CTCAGCTGTGCTGTGCAAGGCTTGAAGTGTCTGACAGTGTCTCAAGGCTCATATGC 1080
Qy 2406 TGACAGCTCTTCAAGTCTGTGAGCTCTGCAACAGCATCCACATGCGGGCAGATGCA 2465

Db	1081	TGACAGCTCTTACAGACTCCTTGAGTCCCTCGCACCAACAGATCCACACTGCCGGCAGAGTGC	1149
QY	2466	GCGCCGTGTCTCTCTGTGTGTTCCTGCTGCTGCCAGGGCTTCTCTCCG	2514
Db	1141	GCGCCGTGTCTCTCTGTGTGTTCCTGCTGCTGCCAGGGCTTCTCTCCG	1189
RESULT 3			
BX356089			
LOCUS			
DEFINITION	BX356089 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI008YG07 5-PRIME, mRNA sequence.		
ACCESSION	BX356089		
VERSION	BX356089.2		
KEYWORDS	EST.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria (to 936)		
AUTHORS	Ll,w.B., Gruber,C., Jeesee,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization Unpublished (2001)		
COMMENT	On May 5, 2003 this sequence version replaced gi:30372027. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7663.f For more information about this cluster, see http://www.genoscope.cns.fr/cdnatse=CSODI008AD04QP1&c=7663.f .		
FEATURES			
Source			
	1..936		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CSODI008YG07"		
	/tissue_type="PLACENTA COT 25-NORMALIZED"		
	/clone_1lb="Homo sapiens PLACENTA COT 25-NORMALIZED"		
	/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."		
ORIGIN			
Query Match	34.5%; Score 923.6; DB 5; Length 936;		
Best Local Similarity	99.8%; Pred.No.3.7e-154;		
Matches 934; Conservative	1; Mismatches 0; Indels 1; Gaps 1;		
QY	1326	GGAGGAGACACCACGCTGGAGATCTTTAAGACCATCTGACGATTTCCGCCCAAG	1385
Db	1	GGAGGAGACACCACGCTGGAGATCTTTAAGACCATCTGACGATTTCCGCCCAAG	60
QY	1386	TTCTGGAACCTGTACCGAATGGAGAAGGAGGCTTTCTAGCATCTTCTGCTACATGAC	1445
Db	61	TTCTGGAACCTGTACCGAATGGAGAAGGAGGCTTTCTAGCATCTTCTGCTACATGAC	120
QY	1446	CACCTACTGAGACACCCAGAGCCCTTACCAAAGAAGATGTGTCTCTGACAGGGCT	1505
Db	121	CACCTACTGAGACACCCAGAGCCCTTACCAAAGAAGATGTGTCTCTGACAGGGCT	180
QY	1506	GTGCCAAGACAAGTGGGTGTGTAACCATGAGACCTGTGTGGATGAGACATTGGA	1565
Db	181	GTGCCAAGACAAGTGGGTGTGTAACCATGAGACCTGTGTGGATGAGACATTGGA	240
QY	1566	CGGGGAACTTTGGGGAAGTGTTCAGGGAGCGCTTGGAGCCGACACAACCTGTGTGCG	1623
Db	241	CGGGGAACTTTGGGGAAGTGTTCAGGGAGCGCTTGGAGCCGACACAACCTGTGTGCG	300

OY	1626	GTGAAGTCTTGTGCAGAGAGAGCTTCCCACTGACCTTCAAGGCGCAAGTTTCTACAGAAAGCG	1683
Db	301	GTGAAGTCTTGTGCAGAGAGAGCTTCCCACTGACCTTCAAGGCGCAAGTTTCTACAGAAAGCG	360
OY	1686	AGATTCCTGAAGCAGTACAGACCAACCCCAATCGTGGGTCTATTGGTGTGTGCACCCAG	1745
Db	361	AGATTCCTGAAGCAGTACAGACCAACCCCAATCGTGGGTCTATTGGTGTGTGCACCCAG	420
OY	1746	AAGCAGCCCATCTACATCGTCATGAGCTTGTGCAGGGGGGCGACTTTCGACCTTCTC	1805
Db	421	AAGCAGCCCATCTACATCGTCATGAGCTTGTGCAGGGGGGCGACTTTCGACCTTCTC	480
OY	1806	CGCAGGAGGGGGGGCGCGCTGCGGGGTGAAGACTGTGTCGACAGATGATGGGGGATGACACT	1865
Db	481	CGCAGGAGGGGGGGCGCGCTGCGGGGTGAAGACTGTGTCGACAGATGATGGGGGATGACACT	540
OY	1866	GCTTGCAATGAGTACTCTGAGAGCAGCAGTGTGATCCACCGGACCTTGACTCTCGAAC	1925
Db	541	GCTTGCAATGAGTACTCTGAGAGCAGCAGTGTGATCCACCGGACCTTGACTCTCGAAC	600
OY	1926	TGCTCGTGAACAGAGAAATGTCTGAAATCAGTACCTTTGGGATGTCCCGAGAGAA	1985
Db	601	TGCTCGTGAACAGAGAAATGTCTGAAATCAGTACCTTTGGGATGTCCCGAGAGAA	660
OY	1986	GCGCATGGGGGTCTATGAGAGCTCAGAGGGGGGCTCAGACAGTCCCGTGAAGTGAACGCA	2045
Db	661	GCGCATGGGGGTCTATGAGAGCTCAGAGGGGGGCTCAGACAGTCCCGTGAAGTGAACGCA	720
OY	2046	CCTGAGGCCCTTAACTACAGGCGGCTACTCTCCGAAAGCAGCGTGTGAAGCTTTGGCATC	2105
Db	721	CCTGAGGCCCTTAACTACAGGCGGCTACTCTCCGAAAGCAGCGTGTGAAGCTTTGGCATC	780
OY	2106	TTGCTCTGGGAGAGACTTCAAGCTTGGGGGCTCTCCCTATTCCTCAACCTCAGCAATCAGCA	2164
Db	781	TTGCTCTGGGAGAGACTTCAAGCTTGGGGGCTCTCCCTATTCCTCAACCTCAGCAATCAGCA	840
OY	2165	GACACGGGAGTGTGTGAGAGAGGGGGGCGCTGCTCCCTGCCAGAGCTGTGTCTGTATGC	2224
Db	841	GACACGGGAGTGTGTGAGAGAGGGGGGCGCTGCTCCCTGCCAGAGCTGTGTCTGTATGC	900
OY	2225	CGTGTTCAGGCTCATGAGCAGCAGTGTGCGGCTATGA	2260
Db	901	CGTGTTCAGGCTCATGAGCAGTGTGCGGCTATGA	936

RESULT 4
BX356088/c

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BX356088 980 bp mRNA linear EST 08-APR-2004
BX356088 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSOD1008YG07 3-PRIME, mRNA sequence.
BX356088
BX356088.2 GI:46288403
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 980)
L.I.W.B., Gruber,C., Jeejee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30369992.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7663.F

For more information about this cluster, see
<http://www.genoscope.cne.fr/cdna?cs=CSODI008AD04NP1ec=7663.f>.

FEATURES

source

Location/Qualifiers
 1. 380
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI008Y007"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 34.0%; Score 908.8; DB 5; Length 980;
 Best Local Similarity 95.2%; Pred. No. 1.6e-151;
 Matches 926; Conservative 24; Mismatches 21; Indels 2; Gaps 2;

1533 CATGAGAGCTGGTGTGGTGAGCAGATTGACGCGGGAACTTTGGCGAAGTGTTCAGC 1592
 980 CATGAGAGCTGGTG-TGGGTGAGCAGATTGACGCGGGAACTTTGGCGAAGTGTTCAGC 922

1593 GGAACGCTCGAGCGGACAAACCCCTGGTGGCGGTGAAGCTTTGTGAGAGAGCGTCCCA 1652
 921 GGAACGCTCGAGCGGACAAACCCCTGGTGGCGGTGAAGCTTTGTGAGAGAGCGTCCCA 862

1653 CCTGACCTCA-GGCGAAGTTTCTACAGAGAGGAGATCTGGAACAGTACAGCCACC 1711
 861 CCTGACCTCAAGGCGCAAGTTTCTACAGAGAGGAGATCTGGAACAGTACAGCCACC 802

1712 CAACATCGTGGCTGTCTATTGGTGTCTGCAACCCAGAGAGCCCATCTACATCGTCA 1771
 801 CAACATCGTGGCTGTCTATTGGTGTCTGCAACCCAGAGAGCCCATCTACATCGTCA 742

1772 GCTGTGCAAGGGGGGAGCTTCTGACCTTCCCTCGGACAGAGGGGGCCGCTGGGGGT 1831
 741 GCTGTGCAAGGGGGGAGCTTCTGACCTTCCCTCGGACAGAGGGGGCCGCTGGGGGT 682

1832 GAAGACTCTGCTGCAATGATGGTGGGATGACAGCTGTCGATGAGTACCTGAGAGCAA 1891
 681 GAAGACTCTGCTGCAATGATGGTGGGATGACAGCTGTCGATGAGTACCTGAGAGCAA 622

1892 GTGTGATCAACCGGAGCTGTGCTGCGAACTGCTGTGAGAGAGAGATGTCTT 1951
 621 GTGTGATCAACCGGAGCTGTGCTGCGAACTGCTGTGAGAGAGAGATGTCTT 562

1952 GAAGACTCTGCTGCAATGATGGTGGGATGACAGCGGATGGGTCTTATGACGCTCAGG 2011
 561 GAAGACTCTGCTGCAATGATGGTGGGATGACAGCGGATGGGTCTTATGACGCTCAGG 502

2012 GGGGCTCAGACAGTCCCGTGAAGTGAACCGGACCTGAGGCGCTTAACTACGCGCGCTA 2071
 501 GGGGCTCAGACAGTCCCGTGAAGTGAACCGGACCTGAGGCGCTTAACTACGCGCGCTA 442

2072 CTCCTCGAAAGGAGCTGTGAGCTTTGGCATCTTGTCTTGAGAGACCTTCAGCTGGG 2131
 441 CTCCTCGAAAGGAGCTGTGAGCTTTGGCATCTTGTCTTGAGAGACCTTCAGCTGGG 382

2132 GGGCTTCCCTTATCCCACTCAGCAATCAGCAACAGGGAATTTGTGAGAGAGGGGG 2191
 381 GGGCTTCCCTTATCCCACTCAGCAATCAGCAACAGGGAATTTGTGAGAGAGGGGG 322

2192 CGCTTGCCTGCGCCARAGCTGTGTCTTATGCTGCTGCTTCAAGCTTATGAGAGAGTCTG 2251
 321 CGCTTGCCTGCGCCARAGCTGTGTCTTATGCTGCTGCTTCAAGCTTATGAGAGAGTCTG 262

2252 GGGCTTATGAGCTTGGCAGCGCGCCAGCTTACAGACATCTTACAGAGAGTGCAGCAT 2311
 261 GGGCTTATGAGCTTGGCAGCGCGCGCCAGCTTACAGACATCTTATACAGAGAGTGCAGCAT 202

2312 CCGAAAGCGGATCGGTGAGGCTGGAGACCCCTTCTCAAGCTGTGTGAGCTTGTGAGAGCT 2371

DB 201 CCGAAAGCGGATCGGTGAGGCTGGAGACCCCTTTTAAAGCTGTGTGCTTGTGAGAGCTT 142

QY 2372 AGGTGAGCTCTCTCAGCGGCTCAGCTCATATGCTGACAGCTCTTTCACAGTCTTGACTC 2431
 141 AGGTGAGCTCTCTCAGCGGCTCAGCTCATATGCTGACAGCTCTTTCACAGTCTTGAGAGCT 82

QY 2432 CTGCCACCGATCTCCACTGCGGACAGATGACAGCGCGGTCTCTGTGTGCTGCTG 2491
 81 CTGCCACCGATCTCCACTGCGGACAGATGACAGCGCGGTCTCTGTGTGCTGCTGCTG 22

QY 2492 TGTGTCAGAGGCT 2504
 21 NTATGCCAAGCTT 9

DB

RESULT 5
 AK004587
 LOCUS
 DEFINITION
 Mus musculus adult male lung cDNA, RIKEN full-length enriched library, clone:1200003015 product:similar to PROTO-ONCOGENE TYROSINE-PROTEIN KINASE FES/FPS (BC 2.7.1.112) (C-FES) [Mus musculus], full insert sequence.

ACCESSION
 AK004587.1 GI:12835858
 VERSION
 AK004587
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Taahito, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20530913
 11076861

REFERENCE
 4
 THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE FANTOM CONSORTIUM.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 5
 THE FANTOM CONSORTIUM AND THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE I & II TEAM.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

JOURNAL
 REFERENCE
 AUTHORS
 6 (bases 1 to 3373)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arawaka, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,

Db 1429 CAGAGCCCAAGCGCTGCTCCCTCTTACAGATGACCGCCATTCACCTCTCCACGAGCA 1488

Qy 1313 GAGCGAGAGGGGGAGAGACCCACGCTGAGATCTTTAAGAGCCACATCTCAGGAAT 1372

Db 1489 GAGCGAGAGGGGGAGAGACCCACGCTGAGATCTTTAAGAGCCACATCTCAGGAAT 1548

Qy 1373 CTTCCGCCCCAAGTTCTGACACTGTACCGACTGGA 1408

Db 1549 CTTCCGCCCCAAGTTCTGACACTGTACCGACTGGA 1584

RESULT 6
ALS42976 936 bp mRNA linear EST 24-MAR-2004
LOCUS ALS42976 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YP02
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION ALS42976
VERSION ALS42976.3 GI:45718543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 936)
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:30548660.
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7663.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DE013DH01QP1c=7663.f.
FEATURES
Source 1..936
Location/Qualifiers
1..936
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE013YP02"
/issue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and EcoRV sites of the
pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 33.1%; Score 885.4; DB 1; Length 936;
Best Local Similarity 98.8%; Pred. No. 2.4e-147;
Matches 900; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

Qy 29 TCCGAGGCGCTGCCAGACAGCTGCCGCGAACACACATATAGGCTTCTCTTCA 88

Db 14 TCCCGGAGATGTCACAGACAGCTGCCGCGAACACACATATAGGCTTCTCTTCA 73

Qy 89 GCTGTGACAGCCCGCAGGCGACCGGGTCTCTGACGAATATGACGAGCCGAGCTTGTCT 148

Db 74 GCTGTGACAGCCCGCAGGCGACCGGGTCTCTGACGAATATGACGAGCCGAGCTTGTCT 133

Qy 149 ACTGAGGCGATGAGAAATGATGAGCGCGATCAAGAGTGAAGAGGAGATGATGACG 208

Db 134 ACTGAGGCGATGAGAAATGATGAGCGCGATCAAGAGTGAAGAGGAGATGATGACG 193

Qy 209 ACTGCTTCAACACATCTCTCTGACAGACAGTGGGGGCCAGAGCCGGCCATCAGCCCTGA 268

Db 194 ACTGCTTCAACACATCTCTCTGACAGACAGTGGGGGCCAGAGCCGGCCATCAGCCCTGA 253

Qy 269 CAGCCCATCAGTCACTGCTGAGGCTGAGATCAACAGCCAACTGAGGGCTTGAAGCCCTT 328

Db 254 CAGCCCATCAGTCACTGCTGAGGCTGAGATCAACAGCCAACTGAGGGCTTGAAGCCCTT 313

Qy 329 GCTGGGGCAGACCGAGAGATCTTGAATCAAGGGCCCTGAGCAAGCTTGAAGCTTGTCT 388

Db 314 GCTGGGGCAGACCGAGAGATCTTGAATCAAGGGCCCTGAGCAAGCTTGAAGCTTGTCT 373

Qy 389 CCGGAAACGAGCAGACTTCGCAAGACTTACAGCAGAGCTGGCAGAGCTCAGAGGA 448

Db 374 CCGGAAACGAGCAGACTTCGCAAGACTTACAGCAGAGCTGGCAGAGCTCAGAGGA 433

Qy 449 GCTCAACAAAGCCACAGCAGACATTTGAAGAGCTGAAAGAGCTTGAAGCTTGTCT 508

Db 434 GCTCAACAAAGCCACAGCAGACATTTGAAGAGCTGAAAGAGCTTGAAGCTTGTCT 493

Qy 509 ACGGACAGTGGCCCAAGCCCAAGCTTCAAGAGAGCCCAAGAGAGAGAGAGAGAGAG 568

Db 494 ACGGACAGTGGCCCAAGCCCAAGCTTCAAGAGAGCCCAAGAGAGAGAGAGAGAGAG 553

Qy 569 CAAGGCAAGAGCAAGATATGTCGAGAGCTTGAAGCTTGTCTCAACCAAGAGCT 628

Db 554 CAAGGCAAGAGCAAGATATGTCGAGAGCTTGAAGCTTGTCTCAACCAAGAGCT 612

Qy 629 TGTGCTGGGGCGTGGGGCTGCGCAGCTACACAGCAGCAGCAGCAGCTTGTCT 688

Db 613 TGTGCTGGGGCGTGGGGCTGCGCAGCTACACAGCAGCAGCAGCAGCTTGTCT 672

Qy 689 CCGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748

Db 673 CCGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732

Qy 749 CCGTCAAGAAATCTGAGATTTAGCAGCTTGTGAGATGAGTGGTGGCTTCAAGC 808

Db 733 CCGTCAAGAAATCTGAGATTTAGCAGCTTGTGAGATGAGTGGTGGCTTCAAGC 792

Qy 809 GAGATGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868

Db 793 GAGATGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852

Qy 869 AAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928

Db 853 AAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912

Qy 929 GAGGGGTGAC 939

Db 913 GAGGGGTGAC 923

RESULT 7
BM456755 1029 bp mRNA linear EST 05-FEB-2002
LOCUS BM456755
DEFINITION AGENCOURT_6403957 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583652
ACCESSION BM456755
VERSION BM456755.1 GI:18505795
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 1029)
NIH-MGC http://mhc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12347 row: b column: 05
 High quality sequence stop: 704.
 Location/Qualifiers

FEATURES

source

1..1029
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5583652"
 /issue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH MGC 93"
 /notes="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dt primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC library."

ORIGIN

Query Match 32.8%; Score 875.8; DB 4; Length 1029;
 Best Local Similarity 98.0%; Pred. No. 1.2e-145;
 Matches 939; Conservative 0; Mismatches 13; Indels 6; Gaps 5;

2 CCGGGGTCGGCAACCGGGCTGAGTCCGTCCGAGCCGTCGAGACAGCTGCCCTGGG 61
 11 CCGGGGTCGGCAACCGGGCTGAGTCCGTCCGAGCCGTCGAGACAGCTGCCCTGGG 70
 62 GAACAGCACTATGGGCTTCTCTCTGAGTGTGACAGCCCGGAGGGGCGAGGGGCTCTGCA 121
 71 GAACAGCACTATGGGCTTCTCTCTGAGTGTGACAGCCCGGAGGGGCGAGGGGCTCTGCA 130
 122 GCAATGTCAGAGAGCGGAGCTTCTCTCTGAGTGTGACAGCCCGGAGGGGCGAGGGGCTCTGCA 181
 131 GCAATGTCAGAGAGCGGAGCTTCTCTCTGAGTGTGACAGCCCGGAGGGGCGAGGGGCTCTGCA 190
 182 GGTCAAGATGACAGAGAGTATGACAGACCTTCAACATGTCCTTCGAGAGCACTGAG 241
 191 GGTCAAGATGACAGAGAGTATGACAGACCTTCAACATGTCCTTCGAGAGCACTGAG 250
 242 GGGCCAGAGCCGGGCGCATGAGCCCTGACAGCCCGCATGAGTCTCTGGGCTTGAATCAC 301
 251 GGGCCAGAGCCGGGCGCATGAGCCCTGACAGCCCGCATGAGTCTCTGGGCTTGAATCAC 310
 302 CAGCCAACTGAGAGGCTTGTGCTGCTGCGGACAGAGCGAGAGATCTGAATCTGAG 361
 311 CAGCCAACTGAGAGGCTTGTGCTGCTGCGGACAGAGCGAGAGATCTGAATCTGAG 370
 362 GGGCCAGAGCCGGGCGCATGAGCCCTGACAGCCCGCATGAGTCTCTGGGCTTGAATCAC 421
 371 GGGCCAGAGCCGGGCGCATGAGCCCTGACAGCCCGCATGAGTCTCTGGGCTTGAATCAC 430
 422 CAGAGTGTGAG 481
 431 CAGAGTGTGAG 490
 482 GCTGAAGAGCCAGTACAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541
 491 GCTGAAGAGCCAGTACAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 550
 542 GAGAGCCAG 601
 551 GAGAGCCAG 610
 602 GAAGCTTTTGTCTACCAACAACCGCTATGTGCTGGGCTGCGGCTGCGGCTGCGGCTGCGG 661
 611 GAAGCTTTTGTCTACCAACAACCGCTATGTGCTGGGCTGCGGCTGCGGCTGCGGCTGCGG 670
 662 CAGAGCAACAACAGAGCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
 671 CAGAGCAACAACAGAGCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 730

QY 722 GGAGATGGCTTGCATCTGTAAGAGATCTGAGAGAACTGAGATTAGAGACTGTGT 781
 DB 731 GGAGATGGCTTGCATCTGTAAGAGATCTGAGAGAACTGAGATTAGAGACTGTGT 790
 QY 782 GCAGAGTGAAGTGTGTGAGGCTTCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 DB 791 GCAGAGTGAAGTGTGTGAGGCTTCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 850
 QY 841 CTGAGAGCTGAGTGAAG 899
 DB 851 CTGAGAGCTGAGTGAAG 909
 QY 900 TGTGTGACG-TTCGATGAGTCACTGCTTGA--GAGAGGTGAACCGCTGAGAGCTGGGG 954
 DB 910 TGTGTGACGTTTCGATGAGTCACTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 967

RESULT 8
 CR622908
 LOCUS
 DEFINITION
 full-length cDNA clone CSDBE004YN09 of Placenta of Homo sapiens
 (human).
 accession
 CR622908.1 GI:50503715
 VERSION
 HTP; CNSLT cDNA.
 KEYWORDS
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 863)
 Li W.B., Gruber C., Jessee J. and Polayes D.
 Full-length cDNA libraries and normalization
 JOURNAL
 Unpublished
 CONTACT : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 REMARK
 Faraday Avenue
 2 (bases 1 to 863)
 Genoscope.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT
 Direct Submission
 Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 Location/Qualifiers

FEATURES
 source
 1..863
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSDBE004YN09"
 /issue_type="Placenta"
 /plasmid="pCMVSPORT_6"

ORIGIN

Query Match 32.3%; Score 863; DB 3; Length 863;
 Best Local Similarity 100.0%; Pred. No. 2.3e-143;
 Matches 863; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1688 AAGATCTGAAGACAGTACAGCAACCCCAATGCGCTCATTTGATGTGACCCGAG 1745
 DB 1 AAGATCTGAAGACAGTACAGCAACCCCAATGCGCTCATTTGATGTGACCCGAG 60
 QY 1746 AAGCAACCCATATACATGATGAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1805
 DB 61 AAGCAACCCATATACATGATGAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 1806 CGCAG 1865
 DB 121 CGCAG 180


```

QY 1866 GCTGGATGAGATGCTGAGAGACAGTGTGATCCACCGGAGCTGTGCTCGGAAC 1925
DB 181 GCTGGATGAGATGCTGAGAGACAGTGTGATCCACCGGAGCTGTGCTCGGAAC 240
QY 1926 TGGCTGCTGACAGAGAGAAATGCTCTGAAGATCAGTGTGATGTCCTCGAGAGAA 1985
DB 241 TGGCTGCTGACAGAGAGAAATGCTCTGAAGATCAGTGTGATGTCCTCGAGAGAA 300
QY 1986 GCCGATGGGAGTCTATGCAAGCTTCAAGGAGGCTTCAAGTCCCGTGAAGTGACCGCA 2045
DB 301 GCCGATGGGAGTCTATGCAAGCTTCAAGGAGGCTTCAAGTCCCGTGAAGTGACCGCA 360
QY 2046 CCTGAGGCGCTTAACTACAGGCGGCTCTCTCCGAAAGGAGCTGTGAGCTTGGCATC 2105
DB 361 CCTGAGGCGCTTAACTACAGGCGGCTCTCTCCGAAAGGAGCTGTGAGCTTGGCATC 420
QY 2106 TTGCTCTGGAGAGACTTCAAGCTTCAAGGAGGCTTCCCGTGAAGTGAATGACGAG 2165
DB 421 TTGCTCTGGAGAGACTTCAAGCTTCAAGGAGGCTTCCCGTGAAGTGAATGACGAG 480
QY 2166 ACAAGGAGTGTGTGAGAGAGAGGAGGCGCTCTGCTGCTCCAGAGCTGTGTCTGATGCC 2225
DB 481 ACAAGGAGTGTGTGAGAGAGAGGAGGCGCTCTGCTGCTCCAGAGCTGTGTCTGATGCC 540
QY 2226 GTTGTCAAGCTCTATGAGAGAGTGTGAGGCTTATGAGCTTGGAGGAGGCGGCTTCAAG 2285
DB 541 GTTGTCAAGCTCTATGAGAGAGTGTGAGGCTTATGAGCTTGGAGGAGGCGGCTTCAAG 600
QY 2286 ACCATCTACAGAGAGCTGAGAGAGCTCCGAAAGGAGGAGTGTGAGGAGTGGAGCCGCTT 2345
DB 601 ACCATCTACAGAGAGCTGAGAGAGCTCCGAAAGGAGGAGTGTGAGGAGTGGAGCCGCTT 660
QY 2346 CTCAGCTGTGTGAGAGAGCTTCAAGGAGCTTCTCAAGGAGCTTCAAGCTCATATGC 2405
DB 661 CTCAGCTGTGTGAGAGAGCTTCAAGGAGCTTCTCAAGGAGCTTCAAGCTCATATGC 720
QY 2406 TGAACAGCTCTTCAAGAGTCTGAGAGCTCTGCTCCAGAGAGTCAATGAGTGGAGTGA 2465
DB 721 TGAACAGCTCTTCAAGAGTCTGAGAGCTCTGCTCCAGAGAGTCAATGAGTGGAGTGA 780
QY 2466 GAGCGCTGTCTCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2525
DB 781 GAGCGCTGTCTCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 2526 TAAACCACTTGTGACCACTGAA 2548
DB 841 TAAACCACTTGTGACCACTGAA 863

```

```

RESULT 9
BX359009/c 989 bp mRNA linear EST 08-APR-2004
LOCUS BX359009 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1051YO20 3-PRIME, mRNA sequence.
ACCESSION BX359009
VERSION BX359009.2 GI:46305682
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 989)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30376296.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned

```

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7663.f

For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?cs=CS0D1051BH1ONP1ec=7663.f>.

FEATURES

Location/Qualifiers

1..989

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1051YO20"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 32.0%; Score 855.2; DB 5; Length 989;

Best Local Similarity 95.8%; Pred. No. 5.6e-142;

Matches 944; Conservative 14; Mismatches 18; Indels 9; Gaps 8;

```

QY 1510 CCAAGGACAGTGGGCTGACGACATGAGAGACCTGTGTGGGTGAGCAGATTGACGGG 1569
DB 989 CCAAGGACAGTGGGCTGACGACATGAGAGACCTGTGTGGGTGAGCAGATTGACGGG 934
QY 1570 GGAACCTTTGGGCGAAGTGTTCAGCGAGCGCTCGAGCGGAGCAACAC-CTGTGTGGCGGTG 1628
DB 933 GGAACCTTTGGGCGAAGTGTTCAGCGAGCGCTCGAGCGGAGCAACACMCKTGTGTGGCGGTG 874
QY 1629 AAGTCTTGTGAGAGAGCGTCCCACTGACCTTCAAGGCGCAAGTTTCAAGAGAGGAGG 1688
DB 873 AAGTCTTGTGAGAGAGCGTCCCACTGACCTTCAAGGCGCAAGTTTCAAGAGAGGAGG 814
QY 1689 ATCTGGAACAGTGAAGCAACCCCAATGAGGCTGATGTGTGTGTGACCCGAGAG 1748
DB 813 ATCTGGAACAGTGAAGCAACCCCAATGAGGCTGATGTGTGTGTGACCCGAGAG 754
QY 1749 CAGCCC-ATCTACATGCTATGAGAGCTTGTGAGAGGAGGCGAGCTTCTGACCTTCTCCG 1807
DB 753 CAGCCCATCTACATGCTATGAGAGCTTGTGAGAGGAGGCGAGCTTCTGACCTTCTCCG 694
QY 1808 CAGGAGGAGGAGGCGGCTGCGGAGTGAAGCTGCTGCAAGATGTGTGGAGATGACGCTGC 1867
DB 693 CAGGAGGAGGAGGCGGCTGCGGAGTGAAGCTGCTGCAAGATGTGTGGAGATGACGCTGC 635
QY 1868 TGGCATGAGATCCTTGAAGAGCAAGTGTGATCCACC-AGGACCTGTGCTCGGAACT 1926
DB 634 TGGCATGAGATCCTTGAAGAGCAAGTGTGATCCACCAGGAGACTGTGCTCGGAACT 575
QY 1927 GCTGTGACAGAGAGAAATGCTCTGAAGATCAGTGAATGAGTGTGAGTGTGAGAGAG-A 1985
DB 574 GCTGTGACAGAGAGAAATGCTCTGAAGATCAGTGAATGAGTGTGAGTGTGAGAGAG-A 515
QY 1986 GCCGATGGGAGTCTATGCAAGCTTCAAGGAGGCTTCAAGCAAGTCCCGTGAAGTGACCGCA 2045
DB 514 GCCGATGGGAGTCTATGCAAGCTTCAAGGAGGCTTCAAGCAAGTCCCGTGAAGTGACCGCA 455
QY 2046 CCTGAGGCGCTTAACTACAGGCGGCTCTCTCCGAAAGGAGCTGTGAGCTTGGCATC 2105
DB 454 CCTGAGGCGCTTAACTACAGGCGGCTCTCTCCGAAAGGAGCTGTGAGCTTGGCATC 395
QY 2106 TTGCTCTGGAGAGACTTCAAGCTTGGAGGCTTCCCGTGAAGTGAATGACGAG 2165
DB 394 TTGCTCTGGAGAGACTTCAAGCTTGGAGGCTTCCCGTGAAGTGAATGACGAG 335
QY 2166 ACAAGGAGTGTGTGAGAGAGAGGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2225
DB 334 ACAAGGAGTGTGTGAGAGAGAGGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 275
QY 2226 GTTGTCAAGCTCTATGAGAGAGTGTGAGGCTTATGAGCTTGGGAGGCGGCGGCACTTCAAG 2285

```

Db 274 GTGTTACAGGCTATGAGCAGTGTCTGGCCATGAGCCTGGGCAAGCGCBACAGCTTCACG 215
Qy 2286 ACCATTACACAGAGAGCTGACAGACATCCGAAAGCGGCATTCGGTAGGCTGGAGACCCCTT 2345
Db 214 ACCATTACACAGAGAGCTGACAGACATCCGAAAGCGGCATTCGGTAGGCTGGAGACCCCTT 155
Qy 2346 CTCACAGCTGTGAGCTCTGTGAGGCTTAGGTGACAGCTCTCAGCGGCTCCAGCTCATATGC 2405
Db 154 CTCACAGCTGTGAGCTCTGTGAGGCTTAGGTGACAGCTCTCAGCGGCTCCAGCTCATATGC 95
Qy 2406 TGACAGCTCTTACAGCTCTGAGCTCTGACACAGCATTCACATGCCGAGATGCA 2465
Db 94 TGACAGCTCTTACAGCTCTGAGCTCTGACACAGCATTCACATGCCGAGATGCA 35
Qy 2466 GCGCCGTGTCTCTGTGTCTGTCTG 2490
Db 34 GCGCCGTGTCTCTGTGTCTGTCTGTCTG 10

RESULT 10
AL540776 861 bp mRNA linear EST 24-MAR-2004
LOCUS AL540776 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE004YN09
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION AL540776
VERSION AL540776.3 GI:45716373
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 861)
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:30544306.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7663.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DE004CG050P1c=7663.f.

FEATURES
source
1..861
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE004YN09"
/issue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 31.3%; Score 836.6; DB 1; Length 861;
Best Local Similarity 99.4%; Pred. No. 1,1e-138;
Matches 858; Conservative 2; Mismatches 1; Indels 2; Gaps 2;
Qy 1686 AGGATCTGAGAGAGTACAGCAACCCCAACATGCTGCTTCACTTGTCTCCACACCA 1745
Db 1 AGGATCTGAGAGAGTACAGCAACCCCAACATGCTGCTTCACTTGTCTCCACACCA 60
Qy 1746 AACGAGCCCATCTACATCTGATGAGCTGTGTCAAGGGGCGCATCTTCTGACCTTCTC 1805

Db 61 AAGCAGCCCATTTATCATGCTATGAGCTTGTGCAAGGGGGGAGATTCTTGACTTCTCTC 120
Qy 1806 CGCAGAGAGGGGGCCCGCTGGGGGTGAAGACTCTGCTGCAATGCTGGGGAGATGAGAGT 1865
Db 121 CGCAGAGAGGGGGCCCGCTGGGGGTGAAGACTCTGCTGCAATGCTGGGGAGATGAGAGT 180
Qy 1866 GCTGAGATGAGATCCTGAGAGCAAGTGTGATTCACACCGGAGACTGTGCTCGGAAC 1925
Db 181 GCTGAGATGAGATCCTGAGAGCAAGTGTGATTCACACCGGAGACTGTGCTCGGAAC 240
Qy 1926 TGCCTGTGACAGAGAGAGATGTCTGAAGATGAGATTTGGAGATGTCCGAGAGAA 1985
Db 241 TGCCTGTGACAGAGAGAGATGTCTGAAGATGAGATTTGGAGATGTCCGAGAGAA 300
Qy 1986 GCGCATGGGGGTATGAGAGCTTCAGGGGGCTTCAGACAAAGTCCCCGTGAAGTGAACCGGA 2045
Db 301 GCGCATGGGGGTATGAGAGCTTCAGGGGGCTTCAGACAAAGTCCCCGTGAAGTGAACCGGA 360
Qy 2046 CCTGAGAGCCCTTAATACAGGCGGCTACTCTCCGAAGGAGAGTGTGAGACTTTGGAGATC 2105
Db 361 CCTGAGAGCCCTTAATACAGGCGGCTACTCTCCGAAGGAGAGTGTGAGACTTTGGAGATC 420
Qy 2106 TTGCTCTGGAGACCTTACAGCTGAGGGGCTTCCCTTAATCCAACTCAGCAATCAGCAG 2165
Db 421 TTGCTCTGGAGACCTTACAGCTGAGGGGCTTCCCTTAATCCAACTCAGCAATCAGCAG 480
Qy 2166 AACGAGAGTTTGTGAGAGAGGGGGGCGCTGCGCTCCAGAGAGCTGTGCTGATGAGC 2225
Db 481 AACGAGAGTTTGTGAGAGAGGGGGGCGCTGCGCTCCAGAGAGCTGTGCTGATGAGC 540
Qy 2226 GTGTTACAGCTATGAGAGAGTGTGAGGCTTAGAGCTGGAGAGCGGCCAGACTTCACG 2285
Db 541 GTGTTACAGCTATGAGAGAGTGTGAGGCTTAGAGCTGGAGAGCGGCCAGACTTCACG 600
Qy 2286 ACCATTACACAGAGAGCTGACAGCATCCGAAAGCGGCATTCGGTAGGCTGGAGACCCCTT 2345
Db 601 ACCATTACACAGAGAGCTGACAGCATCCGAAAGCGGCATTCGGTAGGCTGGAGACCCCTT 659
Qy 2346 CTCACAGCTGTGAGCTCTGTGAGGCTTAGGTGACAGCTCTCAGCGGCTCCAGCTCATATGC 2405
Db 660 CTCACAGCTGTGAGCTCTGTGAGGCTTAGGTGACAGCTCTCAGCGGCTCCAGCTCATATGC 719
Qy 2406 TGACAGCTCTTACAGCTCTGAGCTCTGACACAGCATTCACATGCCGAGATGCA 2465
Db 720 TGACAGCTCTTACAGCTCTGAGCTCTGACACAGCATTCACATGCCGAGATGCA 779
Qy 2466 GCGCCGTGTCTGTGTGTCTGTGTCTGTGTGTGCTGCAAGGGCTTCTTCCGGGCAAAACA 2525
Db 780 GCGCCGTGTCTGTGTGTCTGTGTGTCTGTGTGTGCTGCAAGGGCTTCTTCCGGGCAAAACA 839

RESULT 11
BX359010 968 bp mRNA linear EST 08-APR-2004
LOCUS BX359010 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1051Y020 5-PRIME, mRNA sequence.
ACCESSION BX359010
VERSION BX359010.2 GI:46307488
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 968)
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30378289.


```

QY 1746 AAGCAGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGACCTTCCGATCCTTC 1805
DB 788 AAGCAGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGACCTTCCGATCCTTC 729
QY 1806 CGCAGCAGAGGGGGCCCGCTGCGGGTGAAGAAGCTCTGCTCAAGATGATGAGGAGTGAAGCT 1865
DB 728 CGCAGCAGAGGGGGCCCGCTGCGGGTGAAGAAGTCTGCTCAAGATGATGAGGAGTGAAGCT 669
QY 1866 GCTGAGCATGAGATGATCTGAGAGAGCAAGTGTGATCCACCGGAGACTGGCTGTGCGGAAC 1925
DB 668 GCTGAGCATGAGATGATCTGAGAGAGCAAGTGTGATCCACCGGAGACTGGCTGTGCGGAAC 609
QY 1926 TGCTGTGTAACAGAGAAATGTCCTGAAGATGATGATCTGAGATGTCCTGCGAGAGAA 1985
DB 608 TGCTGTGTAACAGAGAAATGTCCTGAAGATGATGATCTGAGATGTCCTGCGAGAGAA 549
QY 1986 GCGGATGGGGTCTATGACAGCTCAGGGGGCTCAGCAAGTCCCGTGAAGTGAACGCA 2045
DB 548 GCGGATGGGGTCTATGACAGCTCAGGGGGCTCAGCAAGTCCCGTGAAGTGAACGCA 489
QY 2046 CCTGAGGCGCTTAATCAAGCGCGCTACTCTCTCGAAAGGAGAGTGTGAGCTTTGGCATC 2105
DB 488 CCTGAGGCGCTTAATCAAGCGCGCTACTCTCTCGAAAGGAGAGTGTGAGCTTTGGCATC 429
QY 2106 TTGCTGTGAGAGACCTTCAAGCGCTGCGGGGCTCCCGTATCCCAAGCTCAGCAATGACAG 2165
DB 428 TTGCTGTGAGAGACCTTCAAGCGCTGCGGGGCTCCCGTATCCCAAGCTCAGCAATGACAG 369
QY 2166 ACACGAGAGTTTGTGAGAGAGGGGGGCGCTCTGCTCCGCAAGAGCTGTCTGATGAGC 2225
DB 368 ACACGAGAGTTTGTGAGAGAGGGGGGCGCTCTGCTCCGCAAGAGCTGTCTGATGAGC 309
QY 2226 GTGTTCAAGCTCAATGAGAGAGTCTGCGCTTATGAGCTTGGGAGCGGCGCCAGCTTCAAG 2285
DB 308 GTGTTCAAGCTCAATGAGAGAGTCTGCGCTTATGAGCTTGGGAGCGGCGCCAGCTTCAAG 249
QY 2286 ACCATCTACAGAGAGTGTGAGAGAGCATCCGAAAGGGGATGAGTGAAGCTGGAGAGCCGCTT 2345
DB 248 ACCATCTACAGAGAGTGTGAGAGAGCATCCGAAAGGGGATGAGTGAAGCTGGAGAGCCGCTT 189
QY 2346 CTCAGAGCTGTGAGCTCTGAGAGGCTTATGAGTGAAGCTCTCAGCGGGCTTCAAGCTATATG 2405
DB 188 CTCAGAGCTGTGAGCTCTGAGAGGCTTATGAGTGAAGCTCTCAGCGGGCTTCAAGCTATATG 129
QY 2406 TGAAGAGCTTTCAGAGTCTGAGCTCTGAGCTCCAGAGATTCAGATCTGCGGAGAGTGA 2465
DB 128 TGAAGAGCTTTCAGAGTCTGAGCTCTGAGCTCCAGAGATTCAGATCTGCGGAGAGTGA 69
QY 2466 GGGCGGTGCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2525
DB 68 GGGCGGTGCTCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10
QY 2526 TAAACACAC 2534
DB 9 KAAACAC 1

```

```

RESULT 13
LOCUS BX381726 939 bp mRNA linear EST 26-APR-2004
DEFINITION BX381726 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1069YK21 5-PRIME, mRNA sequence.
ACCESSION BX381726 GI:46570766
VERSION BX381726
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 939)
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization

```

JOURNAL COMMENT

Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:30443719.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7663.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna7663f10p1ec=7663.f.
Location/Qualifiers

FEATURES

source
1..939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1069YK21"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6
vector. Library was normalized."
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 30.1%; Score 804.4; DB 5; Length 939;
Best Local Similarity 97.3%; Pred. No. 6.3e-133;
Matches 876; Conservative 13; Mismatches 4; Indels 7; Gaps 7;

```

QY 41 CGAGAGCAGCTGCGCGCGGAGACAGCACTTGTGGCTTCTTGTGAGCTGTGAGCGCC 100
DB 1 CGAGAGCAGCTGCGCGCGGAGACAGCACTTGTGGCTTCTTGTGAGCTGTGAGCGCC 60
QY 101 CCAGGGCCAGCGGGGCTCTGCAAGCAATGAGAGGCGCAGCTTGTCTTCTGAGGGCAT 160
DB 61 CCAGGGCCAGCGGGGCTCTGCAAGCAATGAGAGGCGCAGCTTGTCTTCTGAGGGCAT 120
QY 161 GAGAAAGTGAATGCGCCGAGCGGGTCAAGATGACAGGAGTATGACAGACTGCTTACCA 220
DB 121 GAGAAAGTGAATGCGCCGAGCGGGTCAAGATGACAGGAGTATGACAGACTGCTTACCA 180
QY 221 CATGTCCTGAGAGCAGTGGGGGCGAGCGGGGCTCAGCGGCTTCAAGCGGCTTCAAG 280
DB 181 CATGTCCTGAGAGCAGTGGGGGCGAGCGGGGCTCAGCGGCTTCAAGCGGCTTCAAG 240
QY 281 TCAGTCTGGGCTGAGATCAGCAGCCAACTGAGGCGCTTCAAGCGGCTTCAAGCGG 340
DB 241 TCAGTCTGGGCTGAGATCAGCAGCCAACTGAGGCGCTTCAAGCGGCTTCAAGCGG 300
QY 341 CGCAGAGATCTGAATCTCAGGCGGCTTCAAGCGGCTTCAAGCGGCTTCAAGCGG 400
DB 301 CGCAGAGATCTGAATCTCAGGCGGCTTCAAGCGGCTTCAAGCGGCTTCAAGCGG 360
QY 401 GAGCTTGTGAGAGCTTCAAGGAGAGTGTGAGAGGCTTCAAGGAGCTTCAAGGAG 460
DB 361 GAGCTTGTGAGAGCTTCAAGGAGAGTGTGAGAGGCTTCAAGGAGCTTCAAGGAG 420
QY 461 CCACAGCCAGCATTTGAGAACTGAGAGCCAGTACAGAGCTTGTGAGAGCGGAGCAGTGC 520
DB 421 CCACAGCCAGCATTTGAGAACTGAGAGCCAGTACAGAGCTTGTGAGAGCGGAGCAGTGC 480
QY 521 CCAGAGCCAGCCAGATGACAGAGGCGCAGCAAGAGCAAGACCTGACAGAGCCAGGA 580
DB 481 CCAGAGCCAGCCAGATGACAGAGGCGCAGCAAGAGCAAGACCTGACAGAGCCAGGA 540
QY 581 CAAATATGAGCCAGAGCTGTGAAGCTTTTCTCAACAACCGGTATGTGTGGGCGT 640
DB 541 CAAATATGAGCCAGCTGTGAAGCTTTTCTCAACAACCGGTATGTGTGGGCGT 599
QY 641 GGGGCTGTGAGCTTCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 699

```

Accession	Sequence	Position
D8	GGGGGCTTGCGCA - STACACCAACAGACACCAACACCACTCTCTGCGCCGGACCTGCTGC	658
QY	GGTCACTGCAGAGACCTGTCACGAGAGATGCGCTTGTCATCTGAAAGAGATCTTCACAGAAAT	759
D8	GGTCACTGCAGAGACCTGTCACGAGAGATGCGCTTGTCATCTGAAAGAGATCTTCACAGAGAAAT	718
QY	ACCTGAGATTTAGCAGCCTGTGTGCGAGATGAGTGGTGGCATTCCACGGAGATGAGCTG	819
D8	ACCTGAGATTTAGCAGCCTGTGTGCGAGATGAGTGGTGGCATTCCACGGAGATGAGCTG	778
QY	CAGGTGCTGCCCGGATCCAGGCTGACGGTGAAGTACCAAGGCTCTCTCGAGAACAGATGCGGT	879
D8	CAGGTGCTGCCCGGATCCAGGCTGACGGTGAAGTACCAAGGCTCTCTCGAGAACAGATGCGGT	837
QY	CCGCACTGACGTCCCACTCTGTGTCACTTTCGATGAATCACTGCTTGAGAGGGGTGAAC	939
D8	CCGCACTGACGT - CCAMCCCTGTGTCAAGTTCGATG - RKVAMTGGCTTTAAAGAGGTGAAC	894

RESULT 14	
AL569718/c	
LOCUS	913 bp mRNA linear EST 05-APR-2004
DEFINITION	AL569718 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE013YP020
ACCESSION	AL569718
VERSION	AL569718.3 GI:46235981
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 913)	Id, W. B., Gruber, C., Jessup, J., and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished (2001)	On Feb 16, 2001 this sequence version replaced gi:30606435.

Genoscope - Centre National de Séquençage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqrel@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7663.f
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna/b=CS05D01DIHDH01NP1&c=7663.f>.

```

/organisms="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00E013YP02"
/clone_type="PLACENTA"
/clone_id="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five primed and enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

	Query Match	Score	DB 1	Length
	Best local similarity	94.5%	Pred. 1.2e-10	
	Matches 850, Conservative 17, Mismatches 27, Indels 5, Gaps 5,			
Oy	1593	GGAGCGCTCGCAGCCGACAAACCTCGTGGCGCGTGAATC-TTGTGAGAGAGCGCTCC	1651	
Dh	913	GGAGCGCTCGCAGCGACCAACACCTCGTGGCGATGAACTTTTGTTCAGAGAGCGCTCC	854	
Oy	1652	ACCTGACCTCAAGGCCAAGTTT-CTACAGGAAGGAGAGATCTTGAGCAGTACAGCCAC	1710	

Db	853	ACCGACATCCAGGCAAGTTTCTACAGGAAGGAGATCTCTGAACGACGACCAACC	794
Oy	1711	CCAACTATGTCGCTCATTTGGTGTCTGCAACCAGAAAGACAGCCATCTAATCTGTCATMG	1770
Db	793	CCAACTATGTCGCTCATTTGGTGTCTGCAACCAGAAAGACAGCCATCTAATCTGTCATMG	734
Oy	1771	AGCTTGTGCAAGGGGGGCGACTTCTGACCTTCTCTCCGCAAGGAGGGGGCCCGCTCGGG	1830
Db	733	AGCTTGTGCAAGGGGGGCGACTTCTGACCTTCTCTCCGCAAGGAGGGGGCCCG - CTGCGGG	675
Oy	1831	TGAAGACTCTGCTGCAATGTGTGGGGGATGCAAGTCTGGCATGTAGTACTGTGAGAGCA	1890
Db	674	TGAAGACTCTGCTGCAATGTGTGGGGGATGCAAGTCTGGCATGTAGTACTGTGAGAGCA	615
Oy	1891	AGTCTGATCCACCGGGACCTGGCTGTGCGAACTGC - CTGGTGACAGAGAAATGTC	1949
Db	614	AGTCTGATCCACCGGGACCTGGCTGTGCGAACTGANTCTGGTAAACARBAATAATTTTC	555
Oy	1950	CTGAAGATCAATGACTTTGGGATGTCCCGAGAGAGAACCGATGGGGTCTATGACGCTCA	2009
Db	554	CTGAAGATCAATGACTTTGGGATTTCCCBAGAGAAACCGATGGGGTCTATGACGCTCA	495
Oy	2010	GGGGGCTCAGACAAATGTCCTCGTGAAGTGAACGACACTGAGGCTTTAACTACAGCCGC	2069
Db	494	GGGGGCTCAGACAAATGTCCTCGTGAAGTGAACGACACTGAGGCTTTAACTACAGCCGC	435
Oy	2070	TACTCTCCGAAACGAGCTGTGAGGCTTTGGGACTTTGGCTCTGGAGAACCTTACAGCTG	2129
Db	434	TACTCTCCGAAACGAGCTGTGAGGCTTTGGGACTTTGGCTCTGGAGAACCTTACAGCTG	375
Oy	2130	GGGGCTCTCCCTATATCCCAACTTCAGACAACTACGACAGACAGGGAGTTTGTGGAAGGGG	2189
Db	374	GGGGCTCTCCCTATATCCCAACTTCAGACAACTACGACAACTACGAGAGTTTGTGGAAGGGG	315
Oy	2190	GGCGGTCTGCCCTTGCCAGAGCTGTGTCTGTATGCCGTGTTCAAGCTTCATGAGCAATGC	2249
Db	314	GGCGGTCTGCCCTTGCCAGAGCTGTGTCTGTATGCCGTGTTCAAGCTTCATGAGCAATGC	255
Oy	2250	TGGGCTCATATGAGCTTGGGCAAGCGGCGCCAGCTTCAGACAACTATACAGAGAGCTGACAGC	2309
Db	254	TGGGCTCATATGAGCTTGGGCAAGCGGCGCCAGCTTCAGACAACTATACAGAGAGCTGACAGC	195
Oy	2310	ATCCGAAGCGGCACTCGGTGAGGCTGGGACCCCTTTCTCAAGCTGTGTGCTCTGCAAGC	2369
Db	194	ATCCGAAGCGGCACTCGGTGAGGCTGGGACCCCTTTAAATAATCGTGTGCAATCTGCAAGC	135
Oy	2370	CTAGTGCAGCTCTCTCAGCGGCTTCAGACTCATATGCTGACAGCTCTTCAAGTCTTGAC	2429
Db	134	CTAGTGCAGCTCTCTCAGCGGCTTCAGACTCATATGCTGACAGCTCTTCAAGTCTTGAC	75
Oy	2430	TCTGTGCAACGACATTCACACTGCGCGGACAGATGACAGCGCGGTCTCTCTGTGTCCC	2488
Db	74	YCCGTGCAACCARMAATCCMACTGCCCGGACAGATGCAAGCCGT - TCTCTACTGTNCCC	17

RESULT	15
LOCUS	BQ708270
DEFINITION	BQ708270 Homo sapiens mRNA linear EST 16-JUL-2002 AGNCOCURT#8484911 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6301278 5'..mRNA sequence.
ACCESSION	BQ708270
VERSION	BQ708270.1 GI:21847169
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 948) NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
JOURNAL	Contact: Robert Strausberg, Ph.D.
COMMENT	

Email: c9abbs-r@mail.nih.gov
 Tissue Procurement: Dr. Mark Watson
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLCM2516 row: o column: 07
 High quality sequence stop: 610.

FEATURES

source

Location/Qualifiers

1. 948
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6301278"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_113"
 /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN

Query Match

Best Local Similarity 28.3%; Score 771.6; DB 5; Length 948;

Matches 818; Conservative 0; Mismatches 34; Indels 6; Gaps 2;

```

QY 294 GAGATACACGCAAACTGAGGCGCTGAGCGGCTGTGCGGACGACGAGAGATCTG 353
DB 1 GAGATACACGCAAACTGAGGCGCTGAGCGGCTGTGCGGACGACGAGAGATCTG 60
QY 354 AACTCAGGGCCCTGAGCAGCTGAGCTGCTCATCCGGAAACGAGCAGCTTCGCAAG 413
DB 61 AACTCAGGGCCCTGAGCAGCTGAGCTGCTCATCCGGAAACGAGCAGCTTCGCAAG 120
QY 414 ACTTACAGCAGCAGTGGCGACGCTGCGACGAGAGCTCACCAAGCCACGACCGAGAC 473
DB 121 ACTTACAGCAGCAGTGGCGACGCTGCGACGAGAGCTCACCAAGCCACGACCGAGAC 180
QY 474 ATTGAGAAGCTGAAGGCGCAGTACCGAGCTCTGSCACGCGGACAGTCCCAAGCCAAAGCGG 533
DB 181 ATTGAGAAGCTGAAGGCGCAGTACCGAGCTCTGSCACGCGGACAGTCCCAAGCCAAAGCGG 240
QY 534 AAGTACACGAGGCGCAGCAAGACCAAGACCTGACAAAGGCCAAGACAAGTATGTGCGG 593
DB 241 AAGTACACGAGGCGCAGCAAGACCAAGACCTGACAAAGGCCAAGACAAGTATGTGCGG 300
QY 594 AGCTGTGGAAGCTCTTGTCTCACCAACCGCTATGTGTGGCGTGGGGCTGGCGCAG 653
DB 301 AGCTGTGGAAGCTCTTGTCTCACCAACCGCTATGTGTGGCGTGGGGCTGGCGCAG 360
QY 654 CTACACACGAGGACCAACGACGCTCTGCTGCGGCGCTGTGCGGCTCATCTGACGAGAC 713
DB 361 CTACACACGAGGACCAACGACGCTCTGCTGCGGCGCTGTGCGGCTCATCTGACGAGAC 420
QY 714 CTGCAAGAGAGATGGCTTGCTGCTGAGAGAGATCTGCAAGATACTGAGATTAGC 773
DB 421 CTGCAAGAGAGATGGCTTGCTGCTGAGAGAGATCTGCAAGATACTGAGATTAGC 480
QY 774 AGCTGTGAGAGATGAGTGTGGCCATTCAACCGGAGATGGCTGCAAGTGTGCTGCCCGC 833
DB 481 AGCTGTGAGAGATGAGTGTGGCCATTCAACCGGAGATGGCTGCAAGTGTGCTGCCCGC 540
QY 834 ATCCAGCTGAGGCTGATGCAAGGCTCTGCGACAGTATGGTCCGACCTGACGTC 893
DB 541 ATCCAGCTGAGGCTGATGCAAGGCTCTGCGACAGTATGGTCCGACCTGACGTC 600
QY 894 CCACCTGTGTCACTTCATGATGATCACTGCTTGAAGAGGATGAACCGCTGAGCCTGGG 953

```

```

DB 601 CCACCTGTGTCACTTCATGATGATCACTGCTTGAAGAGGATGAACCGCTGAGCCTGGG 660
QY 954 GAGCTCCAGCTGAAGAGCTGCTGAGAGCGTGCACACAGCTGACCTCAGTACGAC 1013
DB 661 GAGCTCCAGCTGAAGAGCTGCTGAGAGCGTGCACACAGCTGACCTCAGTACGAC 720
QY 1014 GATGAGCTGCTGTGGCCACCGAGATGCTTTCAGGCGGC--AGAGATGTTTACCGACG 1071
DB 721 GATGAGCTGCTGTGGCCACCGAGATGCTTTCAGGCGGCACGAGAGATGTTTACCGACG 780
QY 1072 TCCAAACAGAGCTCCGAAATGAAGCA---GAACCCACCCCGGAGCGGCTGCAG 1127
DB 781 TCCAAACAGAGCTCCGAAATGAAGAGAGAGAAACCCACCCCGGAGAGCGGCTTG 840
QY 1128 CTGCTGGGCAAGAGGCAA 1145
DB 841 CAGCTGCTTGGGCGCAA 858

```

Search completed: March 19, 2005, 14:10:40
 Job time : 7733.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 01:20:31 ; Search time 1275.71 Seconds
(without alignments)
12408.301 Million cell updates/sec

Title: US-10-660-763-1

Perfect score: 2674
Sequence: 1 tccggggccgcacccgggccc.....aaaaaaaaaaaaaaaaaa 2674

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 295970667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16dec04:*
1: geneeqn19808:*
2: geneeqn19908:*
3: geneeqn20008:*
4: geneeqn2001as:*
5: geneeqn2001as:*
6: geneeqn2002as:*
7: geneeqn2002bs:*
8: geneeqn2003as:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004as:*
13: geneeqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2674	100.0	2674 6	ABA94500 Human pro
2	2674	100.0	2674 8	ABV75413 Human kin
3	2674	100.0	2674 12	ADM28580 Human pro
4	2490.2	93.1	2599 12	ADK71888 Human kin
5	2326.4	87.0	2889 6	AAS95001 Human DNA
6	1953	73.0	2623 13	ADR39816 Human kin
7	1552.2	58.0	2680 12	ADH43089 CRAM prot
8	856	32.0	856 1	AAN70060 feg/fps p
9	375	14.0	2818 11	ADI71052 Gene enco
10	374	14.0	527 13	ADQ54298 Novel can
11	362.4	13.6	1263 6	ABK84514 Human CDN
12	362.4	13.6	15297 6	ABA94501 Human pro
13	362.4	13.6	15297 8	ABV75414 Human kin
14	362.4	13.6	15297 12	ADM28582 Human pro
15	328.8	12.3	2584 11	ADI71054 Gene enco
16	326.8	12.2	449 9	ACH14291 Human adu
17	296.2	11.1	361 6	AAS17052 Human pro
18	296.2	11.1	361 10	AAD61604 Human pro
19	296.2	11.1	361 10	ABX14976 Human pro
20	271.6	10.2	1779 8	ACC57665 Drosophi

21	271.6	10.2	3198 4	ABL18403	Ab118403 Drosophi
22	271.6	10.2	3369 4	ABL14457	Ab114457 Drosophi
23	201.2	7.5	3286 12	ADK71885	Adk71885 Human kin
24	201.2	7.5	3875 6	ABZ35414	Abz35414 Human gen
25	199.6	7.5	5616 10	ADE08824	Ad608824 Novel DNA
26	199.2	7.4	2778 12	ADP28218	Adp28218 Human sec
27	198	7.4	2955 6	ABV78117	Abv78117 Human eph
28	198	7.4	2955 6	ABZ35693	Abz35693 Human eph
29	198	7.4	2955 6	ABX09936	Abx09936 Human eph
30	198	7.4	2955 6	ABL91658	Ab191658 Human pol
31	198	7.4	3370 6	ABL67952	Ab167952 Ovary can
32	198	7.4	3370 8	ABZ34861	Abz34861 Coding se
33	198	7.4	3370 10	ADD18447	Add18447 Human pro
34	198	7.4	3370 10	ADE38440	Ad638440 Human pro
35	198	7.4	3370 12	ADU75094	Ad75094 Marker ge
36	198	7.4	3370 13	ADQ80244	Adq80244 Eph1 CDNA
37	185.4	6.9	3042 6	ABV78118	Abv78118 Human eph
38	185.4	6.9	3042 6	ABZ35694	Abz35694 Human eph
39	185.4	6.9	3042 6	ABX09937	Abx09937 Human eph
40	185.4	6.9	3042 6	ABP91559	Ab191559 Human pol
41	185.4	6.9	3346 13	ACN42099	Acn42099 Human dia
42	185.4	6.9	3319 13	ACN37278	Acn37278 Tumour-as
43	185.4	6.9	3921 6	ABV94231	Abv94231 Breast ca
44	185.4	6.9	3921 8	ABZ34865	Abz34865 Coding se
45	185.4	6.9	3921 10	ADB37439	Adb37439 Human can

ALIGNMENTS

RESULT 1	ABV75413	ABV75413
ABV75413	ABV75413	ABV75413
ID	ABA94500 standard; CDNA; 2674 BP.	
XX	ABA94500;	
AC	09-APR-2002 (first entry)	
XX		
DT	Human proto-oncogene tyrosine kinase encoding CDNA.	
XX		
XX	Proto-oncogene tyrosine kinase; potK; tumour; cytostatic; anti-leukemic;	
KW	gene therapy; protein therapy; vaccine; enzymatic-inhibition; human;	
KW	anti-kinase; sb.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	5'UTR	1..71
FT		/*tag= a
FT	CDS	72..2330
FT		/*tag= b
FT		/product= "proto-oncogene tyrosine kinase"
FT	3'UTR	2331..2674
FT		/*tag= c
XX		
XX	US6340584-B1.	
XX		
XX	22-JAN-2002.	
XX		
XX	27-MAR-2001; 2001US-00817180.	
XX		
XX	27-MAR-2001; 2001US-00817180.	
XX		
XX	27-MAR-2001; 2001US-00817180.	
XX		
XX	(PEKE) PE CORP NY.	
XX		
XX	Gan W, Ye J, Di Francesco V, Beasley EM;	
XX		
XX	WPI, 2002-138497/18.	
XX		
XX	F-PSDB; ABB07354.	
XX		
XX	Nucleic acids encoding a proto-oncogene tyrosine kinase, useful for the	
XX	prevention, diagnosis and treatment of e.g. leukemia and lung tumore.	

PS Claim 1, Fig 1A-B; 49pp; English.

XX The invention provides isolated nucleic acid sequences encoding a proto-oncogene tyrosine kinase (POTK). The POTK polynucleotides and protein may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate POTK expression, such as lung and kidney tumors, leukemia and stomach adenocarcinoma. POTK may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of POTK by expressing CC inactive proteins or to supplement the patient's own production of POTK. CC The encoded POTK may be used as an antigen in the production of CC antibodies against POTK and in assays to identify modulators of POTK CC expression and activity. The anti-POTK antibodies and antagonists may be CC used to down regulate expression and activity and as diagnostic agents CC for detecting the presence of POTK in samples. The present sequence CC represents a cDNA encoding the human POTK

SQ Sequence 2674 BP; 662 A; 767 C; 810 G; 435 T; 0 U; 0 Other;

Query Match 100.0%; Score 2674; DB 6; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 2674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCGGGGTCCGACCGGGCTGAGTGGATCCAGGCGGTCCAGAGACGTGCCCTGC 60
DB 1 TCCGGGGTCCGACCGGGCTGAGTGGATCCAGGCGGTCCAGAGACGTGCCCTGC 60
QY 61 GGAAACGACTATGGGCTTCTCTTCTGAGCTGTGACGCCGCCGAGGCTCTCTG 120
DB 61 GGAAACGACTATGGGCTTCTCTTCTGAGCTGTGACGCCGCCGAGGCTCTCTG 120
QY 121 AGCAAAATGACGAGAGCGGAGCTTCTGCTCTA CTGAGGAGCATAGAAATGATGGCCAGC 180
DB 121 AGCAAAATGACGAGAGCGGAGCTTCTGCTCTA CTGAGGAGCATAGAAATGATGGCCAGC 180
QY 181 GGGTCAAGAGTGCACAGGAGATGATGACGACTGCTTACCAACATGTCCTTGACAGACGTG 240
DB 181 GGGTCAAGAGTGCACAGGAGATGATGACGACTGCTTACCAACATGTCCTTGACAGACGTG 240
QY 241 GGGGCGGAGCGGGGCGCATTCAGCCCTGACAGCCCATCTAGTCACTGCTCTGGGCTGAGATCA 300
DB 241 GGGGCGGAGCGGGGCGCATTCAGCCCTGACAGCCCATCTAGTCACTGCTCTGGGCTGAGATCA 300
QY 301 CCAGCCAACTGAGGGCTGAGCGGCTGCTGCTGCGGAGACGCGACAGAGATCTTGAATCTCAG 360
DB 301 CCAGCCAACTGAGGGCTGAGCGGCTGCTGCTGCGGAGACGCGACAGAGATCTTGAATCTCAG 360
QY 361 GGGCCCTGAGCAAGCTGAGCTGCTCATCTCGGGAACGCGACAGCTTGGCAAGACCTTACA 420
DB 361 GGGCCCTGAGCAAGCTGAGCTGCTCATCTCGGGAACGCGACAGCTTGGCAAGACCTTACA 420
QY 421 GCGAGAGTGGCGAGCGCTGACAGAGAGCTCAACAAAGCCCAACACCGAGATTTGAGA 480
DB 421 GCGAGAGTGGCGAGCGCTGACAGAGAGCTCAACAAAGCCCAACACCGAGATTTGAGA 480
QY 481 AGCTGAAGAGCCAGTACCGAGCTCTGCAACGAGACAGTCCCAAGCCCAAGCCCAAGTACC 540
DB 481 AGCTGAAGAGCCAGTACCGAGCTCTGCAACGAGACAGTCCCAAGCCCAAGCCCAAGTACC 540
QY 541 AGGAGCCCAAGCAAGCAAGAGCCGTGACAAAGGCCCAAGCAAGTATGTGCGGAGCTGT 600
DB 541 AGGAGCCCAAGCAAGCAAGAGCCGTGACAAAGGCCCAAGCAAGTATGTGCGGAGCTGT 600
QY 601 GGAAGCTCTTTGCTCAACCAACCGCTATGTGCTGGGCGTGGGCGCTGAGCTTCAACC 660
DB 601 GGAAGCTCTTTGCTCAACCAACCGCTATGTGCTGGGCGTGGGCGCTGAGCTTCAACC 660
QY 661 ACCAGACCAACCAACGCTCTGTGCGGCGCTGCTGCGGCTGCTGAGAGATCTTCAAG 720
DB 661 ACCAGACCAACCAACGCTCTGTGCGGCGCTGCTGCGGCTGCTGAGAGATCTTCAAG 720
QY 721 AGGAGATGGCTTGCATCTTGAAGAGATCTCTGAGAGAAATCACTTGAGATTAGAGCTTGA 780
DB 721 AGGAGATGGCTTGCATCTTGAAGAGATCTCTGAGAGAAATCACTTGAGATTAGAGCTTGA 780

DB 721 AGGAGATGGCTTGCATCTTGAAGAGATCTCTGAGAGAAATCACTTGAGATTAGAGCTTGA 780
QY 781 TGCAGAGATGAGGTGGTGGCATTTACCGGGAGATGGCTGACGTGCCGATCCAGC 840
DB 781 TGCAGAGATGAGGTGGTGGCATTTACCGGGAGATGGCTGACGTGCCGATCCAGC 840
QY 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGTTCGACACTGACCTTCCACCT 900
DB 841 CTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGGTTCGACACTGACCTTCCACCT 900
QY 901 GTGTCACTTGCATGATGATCTGCTTGAAGAGGTGAACCGCTGAGACCTGGGAGCTTC 960
DB 901 GTGTCACTTGCATGATGATCTGCTTGAAGAGGTGAACCGCTGAGACCTGGGAGCTTC 960
QY 961 AGCTGAAGAGCTGACTGTGAGAGCGTGCACACGCTGACCTCACTGATGACAGATGAGC 1020
DB 961 AGCTGAAGAGCTGACTGTGAGAGCGTGCACACGCTGACCTCACTGATGACAGATGAGC 1020
QY 1021 TGGCTGTGGCCACCGAGATGATGTTTCAAGCGGCGAGAGATGTTTACGACTGCAACAG 1080
DB 1021 TGGCTGTGGCCACCGAGATGATGTTTCAAGCGGCGAGAGATGTTTACGACTGCAACAG 1080
QY 1081 AGCTCCGAAATGAGAGAGAACCCACCCCGGAGCGGCTGACGTGCTGGGCAAGA 1140
DB 1081 AGCTCCGAAATGAGAGAGAACCCACCCCGGAGCGGCTGACGTGCTGGGCAAGA 1140
QY 1141 GGCAGATGCTGAAGAAGCACTGACGGGCTGACAGTACGCTGTGACAGGCGCAAGC 1200
DB 1141 GGCAGATGCTGAAGAAGCACTGACGGGCTGACAGTACGCTGTGACAGGCGCAAGC 1200
QY 1201 TGCAGGCGCAGAGAGATGCTGACAGACCAACCTGAGACCTGGGCGCGGACGCC 1260
DB 1201 TGCAGGCGCAGAGAGATGCTGACAGACCAACCTGAGACCTGGGCGCGGACGCC 1260
QY 1261 CGCTGTGCTGCTCTGACAGATGACCGGCACTTCCACTGTGCTCTGAGACAGAGCGAG 1320
DB 1261 CGCTGTGCTGCTCTGACAGATGACCGGCACTTCCACTGTGCTCTGAGACAGAGCGAG 1320
QY 1321 AGGGGGGAAAGGACACCCAGCGCTGAGATCTTAAAGGCCACATCTCAGAAATCTTCGCGC 1380
DB 1321 AGGGGGGAAAGGACACCCAGCGCTGAGATCTTAAAGGCCACATCTCAGAAATCTTCGCGC 1380
QY 1381 CCAAGTCTCGAAGCTGTACCGACTGGAAGGGGAAAGGCTTCTTACATCTCTTGTCTCA 1440
DB 1381 CCAAGTCTCGAAGCTGTACCGACTGGAAGGGGAAAGGCTTCTTACATCTCTTGTCTCA 1440
QY 1441 TCGACCACTTACTGACACCCAGCGCCCTTCAACAAAGAGTGTGTTCTGCA 1500
DB 1441 TCGACCACTTACTGACACCCAGCGCCCTTCAACAAAGAGTGTGTTCTGCA 1500
QY 1501 GGGCTGTGCGCAAGGACAGTGGGTGCTGAACCATGAGGACCTGATGGGTGAGACAGA 1560
DB 1501 GGGCTGTGCGCAAGGACAGTGGGTGCTGAACCATGAGGACCTGATGGGTGAGACAGA 1560
QY 1561 TTGGAACGGGGAACCTTGGCGAAGTGTTCAGCGGACGCTTGGAGCGCAACACCTTGA 1620
DB 1561 TTGGAACGGGGAACCTTGGCGAAGTGTTCAGCGGACGCTTGGAGCGCAACACCTTGA 1620
QY 1621 TGGCGGTGAAGTCTTGTGAGAGAGCGCTCCACTGACCTTCAAGGCGCAAGTGTCTACAG 1680
DB 1621 TGGCGGTGAAGTCTTGTGAGAGAGCGCTCCACTGACCTTCAAGGCGCAAGTGTCTACAG 1680
QY 1681 AAGCAGAGATCTGGAAGAGTACAGGACCCCAACATGCTGGGTCTCATTTGGTGTGCA 1740
DB 1681 AAGCAGAGATCTGGAAGAGTACAGGACCCCAACATGCTGGGTCTCATTTGGTGTGCA 1740
QY 1741 CCCAAGAGAGCCCATCTTACATCTGATAGAGCTTGTGACGGGCGGCGACTTCTTGA 1800
DB 1741 CCCAAGAGAGCCCATCTTACATCTGATAGAGCTTGTGACGGGCGGCGACTTCTTGA 1800
QY 1801 TCTTCGCAACGAGAGGGGCGCGCTTGGGCTGAGAGCTTGTGCTGAGATGTTGGGAGATG 1860
DB 1801 TCTTCGCAACGAGAGGGGCGCGCTTGGGCTGAGAGCTTGTGCTGAGATGTTGGGAGATG 1860

```

QY 1861 CAGCTGCTGGCAATGAGTACTTGGAGAGCAAGTGTGATCCACCGGGAACCTGGCTGCTC 1920
DB 1861 CAGCTGCTGGCAATGAGTACTTGGAGAGCAAGTGTGATCCACCGGGAACCTGGCTGCTC 1920
QY 1921 GGAACTGCTGGTGAAGAGAAATGCTCTGAAGTCAAGTGAATTGGAGATGCCGAG 1980
DB 1921 GGAACTGCTGGTGAAGAGAAATGCTCTGAAGTCAAGTGAATTGGAGATGCCGAG 1980
QY 1981 AGGAAGCCGATGGGCTCTATGACGCTCAGGGGGCTCAGACAAAGTCCCGTGAAGTGA 2040
DB 1981 AGGAAGCCGATGGGCTCTATGACGCTCAGGGGGCTCAGACAAAGTCCCGTGAAGTGA 2040
QY 2041 CCGCACTTGAAGCCCTTAATCTAGGGCCGCTACTCTCCGAAAGCGAGCTGTGAAGCTTTG 2100
DB 2041 CCGCACTTGAAGCCCTTAATCTAGGGCCGCTACTCTCCGAAAGCGAGCTGTGAAGCTTTG 2100
QY 2101 GCATCTTGTCTGGAGAGACCTTCAAGCTGGGGGCTCCCTCTATCCCAACCTCAGCAATC 2160
DB 2101 GCATCTTGTCTGGAGAGACCTTCAAGCTGGGGGCTCCCTCTATCCCAACCTCAGCAATC 2160
QY 2161 AGCAGACACGGGAGTTTGTGAGAGAGGGGGCCGCTGCTGCCCTGACAGAGCTGTCTTG 2220
DB 2161 AGCAGACACGGGAGTTTGTGAGAGAGGGGGCCGCTGCTGCCCTGACAGAGCTGTCTTG 2220
QY 2221 ATGCCCTGTTCAGGCTCATGAGACAGTGTGGGCTCTATGAGCTGGGAGAGCCGACGCT 2280
DB 2221 ATGCCCTGTTCAGGCTCATGAGACAGTGTGGGCTCTATGAGCTGGGAGAGCCGACGCT 2280
QY 2281 TCAGACCATCTTCAAGAGCTCAGAGAGCTCCGAAAGCGGAGTCCGTTGAGGCTGGGACC 2340
DB 2281 TCAGACCATCTTCAAGAGCTCAGAGAGCTCCGAAAGCGGAGTCCGTTGAGGCTGGGACC 2340
QY 2341 CCGTCTCAAGCTGTGGGCTCTGAGAGCTTGAAGTCAAGCTCTCAAGCGCTCAGACTCA 2400
DB 2341 CCGTCTCAAGCTGTGGGCTCTGAGAGCTTGAAGTCAAGCTCTCAAGCGCTCAGACTCA 2400
QY 2401 TATGTCGACAGCTCTTCAAGAGCTCTGAGAGCTCCGAAAGCGGAGTCCGTTGAGGCTGG 2460
DB 2401 TATGTCGACAGCTCTTCAAGAGCTCTGAGAGCTCCGAAAGCGGAGTCCGTTGAGGCTGG 2460
QY 2461 ATGACAGCGCGTGTCTCTCTGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
DB 2461 ATGACAGCGCGTGTCTCTCTGTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
QY 2521 AACCAATTAACCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
DB 2521 AACCAATTAACCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
QY 2581 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2640
DB 2581 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2640
QY 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674
DB 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674

```

```

KM 59.
XX Homo sapiens.
OS Key
XX 5'UTR Location/Qualifiers
XX 1..71 /tag= a
XX CDS 72..2330 /tag= D
XX 3'UTR /product= "protein kinase"
XX 2331..2674 /tag= c
XX
PN MO200277191-A2.
XX
XX 03-OCT-2002.
XX
XX 27-MAR-2002; 2002WC-US009325.
XX
XX 27-MAR-2001; 2001US-00817180.
XX 06-DEC-2001; 2001US-00003295.
XX
XX (PEKE ) PE CORP NY.
XX
XX Gan W, Ye J, Di Francesco V, Beasley EM;
XX WPI, 2003-029927/02.
XX P-PDB; ABB99046.
XX
XX New human kinase protein, useful for treating or diagnosing disorders
XX associated with an absence of, inappropriate, or unwanted expression of
XX the protein, e.g. inflammation or cancer, in drug screening assays and
XX pharmacogenomics.
XX
XX Claim 4 (b) ; Fig 1; 75pp; English.
XX
XX The invention relates to a newly isolated peptide sequence of a human
XX kinase that is related to the proto-oncogene tyrosine kinase subfamily.
XX The activity of the kinase of the invention may be described as,
XX cytosolic, antiarteriosclerotic, antiinflammatory and antiproliferative.
XX peptides of the invention are useful in assays to determine the
XX biological activity of the protein, in drug screening assays, tissue
XX typing and pharmacogenomic analysis. They are also useful in treating or
XX diagnosing disorders characterised by an absence of, inappropriate, or
XX unwanted expression of the protein, such as inflammation, cancer (e.g.
XX leukemia, lung tumours, kidney tumours or stomach adenocarcinoma),
XX arteriosclerosis, and psoriasis. Nucleic acid molecules of the invention
XX are useful as probes, primers and chemical intermediates in biological
XX assays. The peptide and nucleic acid sequences are useful as models for
XX the development of human therapeutic targets, aid in the identification
XX of therapeutic proteins and serve as targets for the development of human
XX therapeutic agents that modulate kinase protein activity in cells and
XX tissues that express the protein. The protein of the invention may also
XX be useful in gene therapy. The gene encoding the protein of the invention
XX has been localised to human chromosome 15. The current sequence
XX represents the human kinase encoding cDNA
XX
XX Sequence 2674 BP; 662 A; 767 C; 810 G; 435 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 2674; DB 8; Length 2674;
XX Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
XX Matches 2674; Conservative 0; Indels 0; Gaps 0;
XX
XX 1 TCCGGGGTCCGACCGGGCTGAGTGGTCCGAGGCGCTCCGAGAGAGAGTCCCGGTG 60
DB 1 TCCGGGGTCCGACCGGGCTGAGTGGTCCGAGGCGCTCCGAGAGAGAGTCCCGGTG 60
QY 61 GGAAACGACCTATGGGCTTCTCTTCTGAGCTGTGACGCCCCGAGGGCCGCGGCTCTG 120
DB 61 GGAAACGACCTATGGGCTTCTCTTCTGAGCTGTGACGCCCCGAGGGCCGCGGCTCTG 120
QY 121 AGCAATGACGAGGCGGAGCTTCTCTCTGAGGAGGATGAGAAAGTGAATGCGCCAGC 180
DB 121 AGCAATGACGAGGCGGAGCTTCTCTCTGAGGAGGATGAGAAAGTGAATGCGCCAGC 180

```

D	b	121	AGCAAAATGCAGAGAGCCGCAAGCTTGCTCATCTGCAGAGGCAATGAAAGCTGATGGCCCAAGC	180
O	y	181	GGGTCAAGAGTGCACAGGGAGTATGCAGAGACTGTTCACACATATGCTCCTCGAGGACAGTG	240
D	b	181	GGGTCAAGAGTGCACAGGGAGTATGCAGAGACTGTTCACCATATGCTCCTCGAGGACAGTG	240
O	y	241	GGGGCCAGAGCCCGGGCCATTCAGCCCTGACACCCCATATAGTCAATGCTCTGGGGCTGAGATCA	300
D	b	241	GGGGCCAGAGCCCGGGCCATTCAGCCCTGACACCCCATATAGTCAATGCTCTGGGGCTGAGATCA	300
O	y	301	CCAGCCAAACTGAGGGGCTGAGCGGCTGTGTCGGGACAGACGAGAGGATCTGAATCTAG	360
D	b	301	CCAGCCAAACTGAGGGGCTGAGCGGCTGTGTCGGGACAGACGAGAGGATCTGAATCTAG	360
O	y	361	GGCCCTTGAGCAAGCTGAGCCTGTGCATCCGAGAAACGGCAGCAAGCTTGCAAGACTTACA	420
D	b	361	GGCCCTTGAGCAAGCTGAGCCTGTGCATCCGAGAAACGGCAGCAAGCTTGCAAGACTTACA	420
O	y	421	GGAGACATGGCAGCAGCTTGCACAGAGAGCTCACCAACCCACAGCCACAGGACATTTGAGA	480
D	b	421	GGAGACATGGCAGCAGCTTGCACAGAGAGCTCACCAACCCACAGCCACAGGACATTTGAGA	480
O	y	481	AGCTGAAGAGCGACATACGAGCTCTGACGAGAACAGTGCCCAAGCCAAAGGCGCAAGTAC	540
D	b	481	AGCTGAAGAGCGACATACGAGCTCTGACGAGAACAGTGCCCAAGGCGCAAGTAC	540
O	y	541	AGAGAGCCAGCAAAAGACAAAGAACCGGTGACAAAGCCCAAGAACAGATATGTGCGCAGCTGT	600
D	b	541	AGAGAGCCAGCAAAAGACAAAGAACCGGTGACAAAGCCCAAGAACAGATATGTGCGCAGCTGT	600
O	y	601	GGAAAGCTTTTGTGCTCACACACACCGCTATGTGCTGGGGGTGCGGAGCTACACCC	660
D	b	601	GGAAAGCTTTTGTGCTCACACACACCGCTATGTGCTGGGGGTGCGGAGCTACACCC	660
O	y	661	ACCAAGCACACACACCAAGCTCCGCTGCGCCCGGCTGTGCGGATCACTGAGGACCTTGACG	720
D	b	661	ACCAAGCACACACACCAAGCTCCGCTGCGCCCGGCTGTGCGGATCACTGAGGACCTTGACG	720
O	y	721	AGGAGATGGCTTGCATCTCTGAAGAGATCTCTGACGAATACCTTGAGATTAGACCCCTGG	780
D	b	721	AGGAGATGGCTTGCATCTCTGAAGAGATCTCTGACGAATACCTTGAGATTAGACCCCTGG	780
O	y	781	TGCAGGATGAGGTGTGTGGCCATTACCCGGGAGATGGCTGACAGTGTCTGCCGATCCAGC	840
D	b	781	TGCAGGATGAGGTGTGTGGCCATTACCCGGGAGATGGCTGACAGTGTCTGCCGATCCAGC	840
O	y	841	CTGAGGGCTGAGATACCAAGGCTTCCGCGACAGTATGGGTCCGACCTGACCGTCCACACT	900
D	b	841	CTGAGGGCTGAGATACCAAGGCTTCCGCGACAGTATGGGTCCGACCTGACCGTCCACACT	900
O	y	901	GTGTCACTTGATGATGAGTCACTGTCTTGAAGAGGTGAACCGCTGAGGCTGGGAGACTCC	960
D	b	901	GTGTCACTTGATGATGAGTCACTGTCTTGAAGAGGTGAACCGCTGAGGCTGGGAGACTCC	960
O	y	961	AGCTGAACGAGCTGACTGTGAGAGACGTGCACACACGCTGACTCACTGATACAGTGAAC	1020
D	b	961	AGCTGAACGAGCTGACTGTGAGAGACGTGCACACACGCTGACTCACTGATCAGATGAAC	1020
O	y	1021	TGGCTGTGGCCACACGAGATGGTGTTCAGGGCGGAGGAATGGTTATCGAGCTGCACACGG	1080
D	b	1021	TGGCTGTGGCCACACGAGATGGTGTTCAGGGCGGAGGAATGGTTATCGAGCTGCACACGG	1080
O	y	1081	AGCTCCGGAATGAAGAGAGAAACCCACCCCGGGACCGGGTGCAGCTGCTGGGCAAGA	1140
D	b	1081	AGCTCCGGAATGAAGAGAGAAACCCACCCCGGGACCGGGTGCAGCTGCTGGGCAAGA	1140
O	y	1141	GGCAAGTCTGCACAAAGCACTGCAGGGGCTGCAGTATGCGCTGTGCACGACGACAGC	1200
D	b	1141	GGCAAGTCTGCACAAAGCACTGCAGGGGCTGCAGTATGCGCTGTGCACGACGACAGC	1200
O	y	1201	TGCAGGCCACAGAGATTTGCTGCGACACCAAGCTGGAACACTTGGGCCCGGCGACAGCCCC	1260
D	b	1201	TGCAGGCCACAGAGATTTGCTGCGACACCAAGCTGGAACACTTGGGCCCGGCGACAGCCCC	1260

QY	1261	UGCTGTGCTGCTCTCTCTCAGAGATGACCGCCCATCTCAAGTGTCTCTCGAGACAGAGGAG	1320
Db	1261	CGCTGTGCTGCTCTCTCAGAGATGACCGCCATCTCAAGTGTCTCTCGAGACAGAGGAG	1320
QY	1321	AGGGGGAGAGACACCCACGCTGGAGATCTCTTAAGACCAATCTCAGAAATCTTCCGC	1380
Db	1321	AGGGGGAGAGACACCCACGCTGGAGATCTCTTAAGACCAATCTCAGAAATCTTCCGC	1380
QY	1381	CCAAATTCTGAACTCTGACCTGTACCGACTGGAAAGGGAAAGGCTTTCCTAGATTCCTTGTCTCA	1440
Db	1381	CCAAATTCTGAACTCTGACCTGTACCGACTGGAAAGGGAAAGGCTTTCCTAGATTCCTTGTCTCA	1440
QY	1441	TCGACCACTACTGAGACCCAGACGGCCCTCAACCAAGAAAGAGTGTTGTCTTGACCA	1500
Db	1441	TCGACCACTACTGAGACCCAGACGGCCCTCAACCAAGAAAGAGTGTTGTCTTGACCA	1500
QY	1501	GGGCTGTGCCCAAGACAAAGTGGTCTGAAACCATGAGGACCTGTGTTGGTGAGCAGA	1560
Db	1501	GGGCTGTGCCCAAGACAAAGTGGTCTGAAACCATGAGGACCTGTGTTGGTGAGCAGA	1560
QY	1561	TTTGAACGGGGGAATTTTGGGGAAGTGTTCAGCGGAGCGCTGCGAGCGGACAAACCCCTGG	1620
Db	1561	TTTGAACGGGGGAATTTTGGGGAAGTGTTCAGCGGAGCGCTGCGAGCGGACAAACCCCTGG	1620
QY	1621	TGGCGGTGAAGTCTTGTGAGAGACGCTCTCCACCTTGAACCTCAAGGCCCAAGTTTCTACAG	1680
Db	1621	TGGCGGTGAAGTCTTGTGAGAGACGCTCTCCACCTTGAACCTCAAGGCCCAAGTTTCTACAG	1680
QY	1681	AAGGAGAGATCTCTGACACGTAACGCCACCCCAACATGTGCGCTCTATTTGGTGTCTGCA	1740
Db	1681	AAGGAGAGATCTCTGACACGTAACGCCACCCCAACATGTGCGCTCTATTTGGTGTCTGCA	1740
QY	1741	CCCAAGACAGACCCATCTACATCGTCACTGAGAGCTTGTGACAGGGGGGCGACTTCTTGACT	1800
Db	1741	CCCAAGACAGACCCATCTACATCGTCACTGAGAGCTTGTGACAGGGGGGCGACTTCTTGACT	1800
QY	1801	TCTCTCCGACGAGAGGGGGCCCGCTCTGCGGTGAAGACTTGTCTGACAGATGGTGGGGGATG	1860
Db	1801	TCTCTCCGACGAGAGGGGGCCCGCTCTGCGGTGAAGACTTGTCTGACAGATGGTGGGGGATG	1860
QY	1861	CAGCTGTGCGATGAGATGACCTGGAGAGCAAGTGTCTGATCCACCGGACCTGGCTCTC	1920
Db	1861	CAGCTGTGCGATGAGATGACCTGGAGAGCAAGTGTCTGATCCACCGGACCTGGCTCTC	1920
QY	1921	GGAATCTGCTGTGTGACAGAGAAAGATGCTCTGAAGATCAGTGACTTTGGATGTCCGAG	1980
Db	1921	GGAATCTGCTGTGTGACAGAGAAAGATGCTCTGAAGATCAGTGACTTTGGATGTCCGAG	1980
QY	1981	AGGAAGCGAATGGGGGTCTATGACGCTCAAGGGGCTCAAGCAGATCCCGTGAAGTGA	2040
Db	1981	AGGAAGCGAATGGGGGTCTATGACGCTCAAGGGGCTCAAGCAGATCCCGTGAAGTGA	2040
QY	2041	CCGACACTGAGAGCCCTTAACCTACGCGCGCTACTCTCCGAAGCGAGCTGTGAGCTTTG	2100
Db	2041	CCGACACTGAGAGCCCTTAACCTACGCGCGCTACTCTCCGAAGCGAGCTGTGAGCTTTG	2100
QY	2101	GCATCTTGTCTGTGGAGACCTTACGCTGGGGGCTTCCCTTATCCCACTCAGCAATC	2160
Db	2101	GCATCTTGTCTGTGGAGACCTTACGCTGGGGGCTTCCCTTATCCCACTCAGCAATC	2160
QY	2161	AGCAGACACGGGAATTTGTGGAGAAAGGGGGGCGCTGTGCCCTGACAGCTGTGTCTTG	2220
Db	2161	AGCAGACACGGGAATTTGTGGAGAAAGGGGGGCGCTGTGCCCTGACAGCTGTGTCTTG	2220
QY	2221	ATGCGCTGTTCAGGCTCATGAGACGATGTGGGCTTATGAGCTGGGACAGGGGCCACGCT	2280
Db	2221	ATGCGCTGTTCAGGCTCATGAGACGATGTGGGCTTATGAGCTGGGACAGGGGCCACGCT	2280
QY	2281	TCAGCACATCTTACCAAGAGCTGTGAGAGCATCCGAAGCGGACATCCGTTAGGCTTGGAC	2340
Db	2281	TCAGCACATCTTACCAAGAGCTGTGAGAGCATCCGAAGCGGACATCCGTTAGGCTTGGAC	2340


```

QY 1347 ATCTTAAGAGGCACATCTCAGGAATCTTCCGCCCCCAAGTTCTGAACTGTATACCGACTG 1406
DB 1405 ATCTTTAAGAGGCACATCTCAGGAATCTTCCGCCCCCAAGTTCTGAACTGTATACCGACTG 1464
QY 1407 GAAGGGGAAGGCTTCTCTAGCATCTTTTGCTCATGACCACTTACTGAGCACTCCAGCAG 1466
DB 1465 GAAGGGGAAGGCTTCTCTAGCATCTTTTGCTCATGACCACTTACTGAGCACTCCAGCAG 1524
QY 1467 CCCCTTACCAAGAAAGTGTGTGTCTCTGACACAGGGCTGTGTCCCAAGACAAAGTGGGTG 1526
DB 1525 CCCCTTACCAAGAAAGTGTGTGTCTCTGACACAGGGCTGTGTCCCAAGACAAAGTGGGTG 1584
QY 1527 CTGAACCATGAGGACCTGTGTGTGGGTGAGACAGATTGAGACGGGGGAACTTTGGCGAAGTG 1586
DB 1585 CTGAACCATGAGGACCTGTGTGTGGGTGAGACAGATTGAGACGGGGGAACTTTGGCGAAGTG 1644
QY 1587 TTCAAGCGAGCGCTTGCAGACCGACACACCTGTGTGGCGGTGAAAGTCTTGTGAGAGACG 1646
DB 1645 TTCAAGCGAGCGCTTGCAGACCGACACACCTGTGTGGCGGTGAAAGTCTTGTGAGAGACG 1704
QY 1647 CTCCCACTTGACTTCAAGGCGCAAGTTTCTTACAGAAAGCGAGATCCTTGAAGCATACAGC 1706
DB 1705 CTCCCACTTGACTTCAAGGCGCAAGTTTCTTACAGAAAGCGAGATCCTTGAAGCATACAGC 1764
QY 1707 CACCCCAACATGATGAGGTCTCATTTGGTGTCTGACCCAGAAAGGACCAATCTACATCGTC 1766
DB 1765 CACCCCAACATGATGAGGTCTCATTTGGTGTCTGACCCAGAAAGGACCAATCTACATCGTC 1824
QY 1767 ATGAGACTTGTGAGAGGGGGGCGACTTCTGACCTTCTCCGACCGAGGGGGCGCGCTG 1826
DB 1825 ATGAGACTTGTGAGAGGGGGGCGACTTCTGACCTTCTCCGACCGAGGGGGCGCGCTG 1884
QY 1827 CCGGTAAAGACTTGTGACAGATGTGGGGGAGATGACAGCTGTGGCATGTGAGTACTTGGAG 1886
DB 1885 CCGGTAAAGACTTGTGACAGATGTGGGGGAGATGACAGCTGTGGCATGTGAGTACTTGGAG 1944
QY 1887 AGCAAGTGTGATCATCCAGCGGAGACTGTGCTCGGAACTGCTGGTGAACAGAGAAAT 1946
DB 1945 AGCAAGTGTGATCATCCAGCGGAGACTGTGCTCGGAACTGCTGGTGAACAGAGAAAT 2004
QY 1947 GTCTTAAGATCAGTACTTTGGGATGTCCCGAAGAGAAAGCGATGGGGTCTATGACGCC 2006
DB 2005 GTCTTAAGATCAGTACTTTGGGATGTCCCGAAGAGAAAGCGATGGGGTCTATGACGCC 2064
QY 2007 TCAAGGGGGCTCAGACAGTCCCGTGAATGACCGGACCTTGAAGCCCTTAATCTACGCG 2066
DB 2065 TCAAGGGGGCTCAGACAGTCCCGTGAATGACCGGACCTTGAAGCCCTTAATCTACGCG 2124
QY 2067 CGCTACTCCCTCCGAAAGCGAGTGTGAGAGCTTGGCATTTGCTTGGGAGACCTTCAGC 2126
DB 2125 CGCTACTCCCTCCGAAAGCGAGTGTGAGAGCTTGGCATTTGCTTGGGAGACCTTCAGC 2184
QY 2127 CTGAGGGGGCTCCCGTATCCCAACTTACAGAAATCAGACAGACAGGGAAGTTTGTGAGAA 2186
DB 2185 CTGAGGGGGCTCCCGTATCCCAACTTACAGAAATCAGACAGACAGGGAAGTTTGTGAGAA 2244
QY 2187 GGGGGGCTGTGCTGCTGCGGACAGTGTGTCTTGAATGCTGTTCAGGCTCAAGAGACG 2246
DB 2245 GGGGGGCTGTGCTGCTGCGGACAGTGTGTCTTGAATGCTGTTCAGGCTCAAGAGACG 2304
QY 2247 TGTGAGGAGCTATAGCTGTGGGAGCGAGCCAGCTTACAGACATTAACAGAGAGTGCAG 2306
DB 2305 TGTGAGGAGCTATAGCTGTGGGAGCGAGCCAGCTTACAGACATTAACAGAGAGTGCAG 2364
QY 2307 AGCATCCGAAGAGCGGATCGGTGAGAGCGCGCTTCTCAAGCTGTGTGCGCTTGCAG 2366
DB 2365 AGCATCCGAAGAGCGGATCGGTGAGAGCGCGCTTCTCAAGCTGTGTGCGCTTGCAG 2424
QY 2367 GGGCTTGTGAGAGTCTCTACGGGGCTTCACTATATGCTGACAGCTTTCAAGATCTG 2426
DB 2425 GGGCTTGTGAGAGTCTCTACGGGGCTTCACTATATGCTGACAGCTTTCAAGATCTG 2484

```

```

QY 2427 GACTCTGCGCACACATCTCACTGCTCGGCGAGATGACAGCGCGCTGTCTCTGTGTGTC 2486
DB 2485 GACTCTGCGCACACATCTCACTGCTCGGCGAGATGACAGCGCGCTGTCTCTGTGTGTC 2544
QY 2487 CTTGCTGCTGCGCAGGGCTTCTCTTCCGGGCGAGAAACATTAACCATTTGTGCC 2541
DB 2545 CTTGCTGCTGCGCAGGGCTTCTCTTCCGGGCGAGAAACATTAACCATTTGTGCC 2599

RESULT 5
AAS95001
ID AAS95001 standard; DNA; 2889 BP.
XX
AC AAS95001;
XX
DT 14-FEB-2002 (first entry)
XX
DE Human DNA sequence #256 expressed during foam cell differentiation.
XX
KW Human; foam cell differentiation; atherosclerosis; cerebral stroke;
XX cardiovascular disorder; coronary artery disease; gene therapy; de.
XX Homo sapiens.
XX
PN WO200177389-A2.
XX
PD 18-OCT-2001.
XX
PF 04-APR-2001; 2001WO-US011128.
XX
PR 05-APR-2000; 2000US-0195106P.
XX
PA (INCYTE GENOMICS INC.
XX
PI Shifman D, Somogyi R, Lawn R, Selthamer J, Porter GJ, Mikita T;
PI Tai J;
XX
DR WPI; 2002-010925/01.
XX
PT Composition useful for diagnosis of conditions, disorders or diseases
PT associated with atherosclerosis, comprises several polynucleotides that
PT are differentially expressed in foam cell development.
PS
PS Claim 1; Page 298-299; 315pp; English.
XX
CC The present invention relates to the isolation of human polynucleotide
CC sequences that are differentially expressed during foam cell
CC differentiation. The polynucleotide sequences of the invention or a
CC composition comprising these polynucleotides are useful as a high
CC throughput method for detecting altered expression of one or more
CC polynucleotides in a sample. The polynucleotides can be used in the
CC diagnosis of disorders associated with foam cell development such as
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used as
CC PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation
XX
SQ Sequence 2889 BP; 600 A; 873 C; 929 G; 487 T; 0 U; 0 Other;

Query Match 87.0%; Score 2326.4; DB 6; Length 2889;
Best Local Similarity 92.3%; Pred. No. 6; 8e-305;
Matches 2558; Conservative 0; Mismatches 1; Indels 211; Gaps 2;
QY 2 CCGGGGCTCCGACCCGGGCTGAGTGTGCGAGGCGGTGCCAGAGACAGTGCCTGTCG 61
DB 118 CCGGGGCTCCGACCCGGGCTGAGTGTGCGAGGCGGTGCCAGAGACAGTGCCTGTCG 177
QY 62 GAACAGACTATAGGGCTTCTTTCGAGCTGTGACGCCCCCAGGGGCGACGGGGTCTGCA 121
DB 178 GAACAGACTATAGGGCTTCTTTCGAGCTGTGACGCCCCCAGGGGCGACGGGGTCTGCA 237

```



```

Db      2398 ACTCCTCCGAAAGCAGCTGTGGAGCTTGGACTTGTGCTGGAGACCTTCAGCCTG 2457
Qy      2131 GGGCCCTCCCTTATCCCACTCAGCAATCAGAGACAGGGAGTTGAGGAGAGGGGG 2190
Db      2458 GGGCCCTCCCTTATCCCACTCAGCAATCAGAGACAGGGAGTTGAGGAGAGGGGG 2517
Qy      2191 GCGCTGTGCGCCGACAGAGCTGTGTCTGATGCGCTGTTCAAGGCTCATGAGACAGTGTCT 2250
Db      2518 GCGCTGTGCGCCGACAGAGCTGTGTCTGATGCGCTGTTCAAGGCTCATGAGACAGTGTCT 2577
Qy      2251 GGGCCCTATAGCGCTGGGACAGCGGCCAGCTTCAGACCACTTACAGAGCTGACAGACA 2310
Db      2578 GGGCCCTATAGCGCTGGGACAGCGGCCAGCTTCAGACCACTTACAGAGCTGACAGACA 2637
Qy      2311 TCCGAAAGGCGCATGCGTGTGAGAGCTGCGGACCCCTTCTCAAGCTGTGGCTGTGACGGCC 2370
Db      2638 TCCGAAAGGCGCATGCGTGTGAGAGCTGCGGACCCCTTCTCAAGCTGTGGCTGTGACGGCC 2697
Qy      2371 TAGGTGACAGCTCCTCAGCGGCTTCAGACTCATATGCTGACAGCTTCACAGTCTTGACT 2430
Db      2698 TAGGTGACAGCTCCTCAGCGGCTTCAGACTCATATGCTGACAGCTTCACAGTCTTGACT 2757
Qy      2431 CCTGCCACAGCATCACACTGCGGACAGATGACAGCGCGCTGTCTCTGTGTCTCTG 2490
Db      2758 CCTGCCACAGCATCACACTGCGGACAGATGACAGCGCGCTGTCTCTGTGTCTCTG 2817
Qy      2491 CTGCTGCCAGGGGCTTCTCTCCGGGACAAATTAACAACCTTGTGCCCATGAGAAA 2550
Db      2818 CTGCTGCCAGGGGCTTCTCTCCGGGACAAATTAACAACCTTGTGCCCATGAGAAA 2877
Qy      2551 AAAAAAAAAA 2560
Db      2878 AAAAAAAAAA 2887

```

RESULT 6

ADRS9816
ID ADR39816 standard; cDNA; 2623 BP.

AC ADR39816;

DT 18-NOV-2004 (first entry)

DE Human kinase and phosphatase KRP-43 encoding cDNA SEQ ID NO:89.

XX human, kinase and phosphatase protein; KRP, enzyme; cytosolic;
XX antiarteriosclerotic; anticonvulsant; neuroprotective;
XX cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory;
XX chymotrypsin; gene therapy; cell proliferative disorder; cancer;
XX atherosclerosis; neurological disorder; epilepsy; Huntington's disease;
XX stroke; immune disorder; inflammatory disorder; AIDS; allergy;
XX developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
XX KRP-43; single nucleotide polymorphism; SNP; gene; 89.

OS Homo sapiens.

HH Key Location/Qualifiers

FT CDS 95..2389

FT variation /tag= a /product= "kinase and phosphatase KRP-43"

FT /tag= b /replac= (109,C)

XX /standard_name= "single nucleotide polymorphism (SNP)"

XX WO2004074453-A2.

PD 02-SEP-2004.

PF 20-FEB-2004; 2004MO-US005092.

PR 20-FEB-2003; 2003US-0449059P.

PR 19-MAR-2003; 2003US-0456932P.

PR 28-MAR-2003; 2003US-0458844P.

```

PR      09-APR-2003; 2003US-0461678P.
PR      17-APR-2003; 2003US-0463937P.
XX      (INCY-) INCYTE CORP.
XX      Ramkumar J, Margulis JP, Swarnaker A, Chawla NK, Tran UK,
PI      Becha SD, Lee SY, Hafalia Adu, Richardson TW, Khare R, Jiang X;
PI      Jackson AA, Yang J, Gorvad AB;
XX      MPI: 2004-635568/61.
DR      P-PSDB; ADR39770.
XX      New human kinases and phosphatases (KRP) for diagnosing, treating and
PT      preventing diseases or conditions associated with aberrant KRP expression
PR      e.g. cancer, acquired immunodeficiency syndrome, epilepsy, or infections.
XX      Claim 5; SEQ ID NO 89; 299pp; English.
XX
XX      The present sequence encodes the human kinase and phosphatase protein
CC      (KRP), designated KRP-43. The human KRP sequences from the present
CC      invention have cytosolic, antiarteriosclerotic, anticonvulsant,
CC      neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic,
CC      anti-inflammatory and chymotrypsin activities, and can be used in gene
CC      therapy. The human KRP proteins and polynucleotides can be used in
CC      diagnosing, treating and preventing diseases or conditions associated
CC      with the decreased expression or overexpression of KRP, such as cell
CC      proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
CC      epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
CC      allergies), or infectious. They can also be used in assessing the effects
CC      of exogenous compounds on the expression of nucleic acid and amino acid
CC      sequences of KRP. The KRP or its fragments are useful in screening
CC      compounds for effectiveness as agonist or antagonist of the polypeptides,
CC      or in altering the expression of the target polynucleotide and compounds
CC      that specifically bind to or modulate the activity of the polypeptide.
XX
XX      Sequence 2623 BP; 541 A; 786 C; 840 G; 456 T; 0 U; 0 Other;
SQ

```

Query Match 73.0%; Score 1953; DB 13; Length 2623;
Best Local Similarity 85.8%; Pred. No. 1,5e-254;
Matches 2360; Conservative 0; Mismatches 5; Indels 384; Gaps 2;

```

Qy      2 CCGGGGTCCGACCGGGCTGAGTGGTCCGAGGCGGTCCAGAGACGCTGCGTGGC 61
Db      25 CCGGGGTCCGACCGGGCTGAGTGGTCCGAGGCGGTCCAGAGACGCTGCGTGGC 84
Qy      62 GAACGACATTAAGGCTTCTTCTGAGCTGTGACGAGCCCGGACAGGGGTCTCTGCA 121
Db      85 GAACGACATTAAGGCTTCTTCTGAGCTGTGACGAGCCCGGACAGGGGTCTCTGCA 144
Qy      122 GCAATGACAGAGGCGGAGCTTCTGTAAGGAGCATAGAAAGTGAATGAGCCAGCG 181
Db      145 GCAATGACAGAGGCGGAGCTTCTGTAAGGAGCATAGAAAGTGAATGAGCCAGCG 204
Qy      182 GGTCAAGAGTGAAGGAGATATGACAGACTGCTTCAACATGTCCCTGAGGACAGTGG 241
Db      205 GGTCAAGAGTGAAGGAGATATGACAGACTGCTTCAACATGTCCCTGAGGACAGTGG 264
Qy      242 GGGCCAGAGCGGGGCTATAGCCCTGACAGGCCCCCATCATCTAGTCTGTGGCTGAGTAC 301
Db      265 GGGCCAGAGCGGGGCTATAGCCCTGACAGGCCCCCATCATCTAGTCTGTGGCTGAGTAC 305
Qy      302 CAGCCAACTGAGGGGCTGAGCGGCTGTGTGGGAGCAGCAGAGAGATCTGAATCAGG 361
Db      306 ----- 305
Qy      362 GCCCTGAGCAAGCTGAGCTGCTCATTCGGGAGACGGCAGAGCTTCGAAAGCCTACAG 421
Db      306 ----- 305
Qy      422 CAGCAGTGGCAGAGCTGACAGAGAGCTCACCAAGACCCACAGCAGGACATTGAGAA 481
Db      306 -----AGACCCACAGCAGGACATTGAGAA 330

```



```

Db      1248  CCAAGCCGTCCTCTTTCAGAGTGAACCCGATTCACCTCTCCAC-----GGAGCA 1301
Qy      1320  GAGGGGGGAAAGACACCAAGCTGAGATCTTTAAGGCCACATCTCAGAAATCTTCGC 1379
Db      1302  GAGGGGGGAAAGACACCAAGCTGAGATCTTTAAGGCCACATCTCAGAAATCTTCGC 1361
Qy      1380  CCCAAGTTCTC----- 1390
Db      1362  CCCAAGTTCTTATCTCTCCACCACTGACCTTCGCGAGGTGCAAGAGCTCTTAT 1421
Qy      1391  ----- 1390
Db      1422  GAGCAGCTGTGATACACGGGGCTATCCCTGGGCAAGGTAGCTGACTTAAGCAC 1481
Qy      1391  ----- 1390
Db      1482  ACTGGGAGCTTCCTGTTGCGGAGAGCCAGGCAAGAGATGATGATCTGTGATG 1541
Qy      1391  -----GAACTGTATCCGACTTGAA 1409
Db      1542  TGGGATGCGCACCCCGACATTCATCATCTTATGATTAAGATTAAGTACGATGAA 1601
Qy      1410  GGGGAAGGCTTCTCTAGCATTCCTTGTCTATGACCACTACTGAGACCCAGAGCC 1469
Db      1602  GGGGATGGCTTCCCGACATACCTCTGTCTATCACTGCTGTCTCTCCAGCACCC 1661
Qy      1470  CTCACCAAGAGAGTGTGTCTGCAAGGGCTGTGTCCCAAGAGCAAGTGGTGTG 1529
Db      1662  CTTCACAAAGAGAGTGTGTCTGTTCAAGGGCGGTGCCCAAGAGCAAGTGGTGTG 1721
Qy      1530  AACCATGAGAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1589
Db      1722  AAGCAGAGAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1781
Qy      1590  AACGGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1649
Db      1782  AAGTGGCGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1841
Qy      1650  CGACCTGACCTCAAGGCCAAGTTTCTACAGAAAGCAGAGATCTTGAAGAGTAC 1709
Db      1842  CCACCTGACCTCAAGGCCAAGTTTCTACAGAAAGCAGAGATCTTGAAGAGTAC 1901
Qy      1710  CCCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1769
Db      1902  CCCAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1961
Qy      1770  GAGCTTGTGAGAGGGGGGAGCTTCTGACCTTCTGCGAGCGAGGGGGCGGCTGCG 1829
Db      1962  GAGCTTGTGAGAGGGGGGAGCTTCTGACCTTCTGCGAGCGAGGGGGCGGCTGCG 2021
Qy      1830  GTGAAGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1889
Db      2022  GTGAAGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2081
Qy      1890  AAGTGTGTATCAACCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1949
Db      2082  AAGTGTGTATCAACCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2141
Qy      1950  CTGAAGATGATGATCTTTGGAGATGTCCCGAGAGAGCGAGTGGGTCTATGAGCT 2009
Db      2142  CTGAAGATGATGATCTTTGGAGATGTCCCGAGAGAGCTGTGTGTGTGTGTGTGT 2201
Qy      2010  GGGGGGCTGAGACAGTCCCGTGAAGTGAACGCACTGAGGCGCTTAACTAAGGCG 2069
Db      2202  AGCGGCTTCAAGACAGTCCCGTGAAGTGAACGCTGAGGCGCTTAACTAAGGCG 2261
Qy      2070  TACTCTCCGAAAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2129
Db      2262  TACTCTCCGAAAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2321
Qy      2130  GGGGCTTCCCGTATCCCAAGCTCAGCAATCAGCAGACGAGGAGTTGTGAGAGGG 2189

```

```

Db      2322  GGGGCTTCAAGCTTATCCCAAGCTTACCAATCAGCAGACAGGGAGTGTGTGAAAAGGG 2381
Qy      2190  GGGGCTTCCCTGCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2249
Db      2382  CATGCTTCTTGTGCGCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2441
Qy      2250  TGGGCTTATGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2309
Db      2442  TGGGCTTATGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2501
Qy      2310  ATCCGAAAGCGGCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2338
Db      2502  ATCCGAAAGCGGCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2530

RESULT 8
AAAT70060
ID      AAAT70060 standard; cDNA; 856 BP.
XX
AC      AAAT70060;
XX
DT      27-AUG-2003 (revised)
DT      25-MAR-2003 (revised)
DT      27-JAN-1991 (first entry)
XX
DE      fes/fps proto-oncogene-related DNA.
XX
KW      Furin; fur gene; tumour diagnosis; ss.
XX
OS      Homo sapiens.
XX
OS      Felle catur.
XX
FH      Key Location/Qualifiers
FT      mat_peptide 3..636
FT      /*tag= a
FT      /product= "fes/fps proto-oncogene protein"
XX
PN      EP246709-A.
XX
PD      25-NOV-1987.
XX
PF      19-MAY-1987; 87EP-00200940.
XX
PR      20-MAY-1986; 86NL-00001271.
XX
PA      (UYRA-) UNIV STICHTING KATHOLIEKE.
XX
PI      Vandeven WJM, Roebroek AJM, Schalken JA;
XX
DR      WPI; 1987-328946/47.
DR      P-PSDB; AAP70055.
XX
PT      Recombinant DNA contg. the fur gene - used for producing furin protein
PT      and antibodies and as a diagnostic aid in the detection of tumours.
XX
PS      Disclosure; Fig 3; 24pp; English.
XX
CC      The fur gene, encoding furin, is in the human and cat genomes directly
CC      upstream of this fes/fps proto-oncogene sequence. Furin is strongly
CC      expressed in specific types of tumours and labelled RNA or DNA probes of
CC      the fur gene and antibodies against furin can be used for diagnostic
CC      purposes. (See also AAAT70061-62 and AAP70056). (Updated on 25-MAR-2003 to
CC      correct PA field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated
CC      on 27-AUG-2003 to correct OS field.)
XX
SQ      Sequence 856 BP; 164 A; 265 C; 259 G; 168 T; 0 U; 0 Other;
XX

Query Match 32.0%; Score 856; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 1,1e-106;
Matches 856; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1693  TGAAGAGTACAGCCCAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1752

```

```

Db      1  TGAAGAGTACAGACCAACCCAAATGTCGTCTATGTCGTCTGACCCAGAAAGCAGC 60
Qy      1753 GCATCTACATCGTCATAGAGCTTGTGACAGGGGGGCACTTCCGACCTTCCCGACGG 1812
Db      61  CCATCTACATCGTCATAGAGCTTGTGACAGGGGGGCACTTCCGACCTTCCCGACGG 120
Qy      1813 AGGGGGGCGCGCTGCGGGTGAAGACTCTGCTGACAGATGTTGGGGATGACAGCTGTCGCA 1872
Db      121  AGGGGGGCGCGCTGCGGGTGAAGACTCTGCTGACAGATGTTGGGGATGACAGCTGTCGCA 180
Qy      1873 TGAAGTACCTGAGAGCAAGTGTGATCAACCGGAGCTTGCTGCTCGAACTGCTG 1932
Db      181  TGAAGTACCTGAGAGCAAGTGTGATCAACCGGAGCTTGCTGCTCGAACTGCTG 240
Qy      1933 TGAAGAGAAAGATGTCGTAAGATGATGATGATGATGATGATGATGATGATGATG 1992
Db      241  TGAAGAGAAAGATGTCGTAAGATGATGATGATGATGATGATGATGATGATGATG 300
Qy      1993 GGGTCTATGACAGCTGAGGGGGGCTGACAGCAAGTCCCGTGAAGTGAACCGGACCTGAG 2052
Db      301  GGGTCTATGACAGCTGAGGGGGGCTGACAGCAAGTCCCGTGAAGTGAACCGGACCTGAG 360
Qy      2053 CCCTTAATAAGCGCCGCTACTCTCTCCGAAAGCAGCTGTGAGCTTTGGCATCTTCTCT 2112
Db      361  CCCTTAATAAGCGCCGCTACTCTCTCCGAAAGCAGCTGTGAGCTTTGGCATCTTCTCT 420
Qy      2113 GGGAGACCTTCAAGCTGAGGGGGGCTCCCGCTATCCCAACCTGACATGACAGACAGGG 2172
Db      421  GGGAGACCTTCAAGCTGAGGGGGGCTCCCGCTATCCCAACCTGACATGACAGACAGGG 480
Qy      2173 AGTTTGTGAGAGAGGGGGGCGCTCTGCTCCCTGACAGAGCTGTGTCTCTGATGCGGTGTA 2232
Db      481  AGTTTGTGAGAGAGGGGGGCGCTCTGCTCCCTGACAGAGCTGTGTCTCTGATGCGGTGTA 540
Qy      2233 GGCTCATGAGACAGTCTGAGGCTTATGAGCTTGGGACAGGGGCCAGCTTACGACCATCT 2292
Db      541  GGCTCATGAGACAGTCTGAGGCTTATGAGCTTGGGACAGGGGCCAGCTTACGACCATCT 600
Qy      2293 ACCAGAGCTGACAGACATCCGAAAGCGGATGAGAGCTGGGAGCCCGCTTCTCAAGC 2352
Db      601  ACCAGAGCTGACAGACATCCGAAAGCGGATGAGAGCTGGGAGCCCGCTTCTCAAGC 660
Qy      2353 TGGTGGCTCTGAGAGCTTATGAGCTTCTGAGGGGCTCCAGCTCATATGCTGACAGC 2412
Db      661  TGGTGGCTCTGAGAGCTTATGAGCTTCTGAGGGGCTCCAGCTCATATGCTGACAGC 720
Qy      2413 TCTTCAAGTCTGAGACTCTGACCAACAGATCAACACTGCGGAGATGACGGCCGT 2472
Db      721  TCTTCAAGTCTGAGACTCTGACCAACAGATCAACACTGCGGAGATGACGGCCGT 780
Qy      2473 GTCTCTCTGTGTCCTGCTGCTGCGAGAGGCTTCTTCCGGGAGAAACAATAAACCC 2532
Db      781  GTCTCTCTGTGTCCTGCTGCTGCGAGAGGCTTCTTCCGGGAGAAACAATAAACCC 840
Qy      2533 ACTTGTGCCACTGAA 2548
Db      841  ACTTGTGCCACTGAA 856

```

RESULT 9
ADL71052
ID ADL71052 standard; DNA; 2818 BP.
XX
XX ADL71052;

XX 20-MAY-2004 (first entry)

DE Gene encoding type II collagen expression promoting protein, seq id 45.

KW Osteopathic; antiinflammatory; antiarthritis; antiarthritic;
KW gene therapy; type II collagen; expression; cartilage disease;
KW osteoarthritis; cartilage defect; rheumatoid arthritis; human; gene; ds.

```

OS Homo sapiens.
XX
PN MO2003087375-A1.
XX
PD 23-OCT-2003.
XX
PF 16-APR-2003; 2003WO-JP004802.
XX
PR 16-APR-2002; 2002JP-0013908.
PR 19-APR-2002; 2002US-0373594P.
XX
PA (ASAH ) ASAH KASEI KK.
XX
PI Matsuda A, Honda G, Muramatsu S;
XX
DR MPI: 2003-845331/78.
DR P-PSDB; ADL71053.
XX
PT New purified protein that promotes type II collagen expression, useful
PT for preventing and treating a cartilage disease, e.g. osteoarthritis,
PT cartilage defect, or rheumatoid arthritis.
XX
PS Claim 4, SEQ ID NO 45; 271pp; English.
XX
CC The invention relates to a purified protein (I) that promotes type II
CC collagen expression. Also disclosed is an isolated polynucleotide and a
CC encoding (I), a recombinant vector comprising the polynucleotide and a
CC gene therapeutic agent comprising the recombinant vector as an active
CC ingredient. The proteins, genes, agents and methods are useful for
CC preventing and treating a cartilage disease, e.g. osteoarthritis,
CC cartilage defect, or rheumatoid arthritis. The current sequence
CC represents a human gene of the invention encoding a protein that promotes
CC type II collagen expression.
XX
SQ Sequence 2818 BP, 909 A; 558 C; 684 G; 667 T; 0 U; 0 Other;
XX
Query Match 14.0%; Score 375; DB 11; Length 2818;
Best Local Similarity 64.8%; Pred. No. 7,4e-42;
Matches 606; Conservative 0; Mismatches 320; Indels 9; Gaps 3;
Qy 1387 TCTCGAATCTGATACGACCTGGAAGGGAGGCTTCTAGCATCTCTTGTCTATGAGAC 1446
Db 1686 TCGATATATCTGATGATGATGATGAGGACCGGGTTTCAAAACATTCCTCCAGCTTATGATC 1745
Qy 1447 ACTTACTGAGCACCAGCAGCCCTCAACCAAGAGATGATGATGATGATGATGATGATG 1506
Db 1746 ACCACTTCAATCAAGACAGTATCAACCAAGATGATGATGATGATGATGATGATG 1805
Qy 1507 TGCCAAAGACAAG--TGGGTGCTGAACCATGAGACCTGTGTGGTGAAGCAGATTG 1563
Db 1806 TCCCAAGAGATGAAGAAATGGTTCTCAATCATGAGATGTTCAATTGGGAGAAATTA 1865
Qy 1564 GACGGGGGAATTTGGCGGAAGTGTCAAGGAGCCTGAGACCGGACCAACCTCTGTG 1623
Db 1866 GCAAGGGGAATTTGGTGAAGGTATMAAG--GCAACATMAAGATMAAACCTCTGTG 1922
Qy 1624 CGGTGAATCTTGTGAGAGACGCTCCACCTGACTGACCTCAAGGCAAGTTTCTACAGAA 1683
Db 1923 CCATTAAACGTGCAAGAGAAAGCTTCTCAGGAATTAATAAATTTCTACAGAA 1982
Qy 1684 CGAGATCTTGAAGCATGACCAACCCCAACATGCTGCTCATGTGCTGCTGACCC 1743
Db 1983 CCAAAATCTGAAGCAATATGATCAATCCAAATATGTTCAAACTGATAGGCGTGTGACAC 2042
Qy 1744 AGAAGAGCCCATCTACATGCTGATGAGCTTGTGAGGGGGGCACTTCTGACCTTCC 1803
Db 2043 AAAAGACGCTGTCTACATCATTAATGAGAACTGTCCAGAGGGGTGATTTTCTGACATTC 2102
Qy 1804 TCGACAGGAGGGGGCGCGCTGCGGGTGAAGACTGTGTCAGATGATGAGGATGACAG 1863
Db 2103 TGAAGAGAGAGAGAGAGAGCTGAAGCTGAAGAGATGATGATGATGATGATGATG 2162
Qy 1864 CTGCTGTCATGAGTACTGAGAGACCAAGTGTGATCCACCGGAGCTGGCTGTGGA 1923

```


Human cDNA differentially expressed in granulocytic cells #1085.

Human; se; granulocytic cell; DNA chip; bacterial infection; viral infection; parasitic infection; protozoal infection; fungal infection; sterile inflammatory disease; psoriasis; rheumatoid arthritis; glomerulonephritis; asthma; thrombosis; cardiac reperfusion injury; renal reperfusion injury; ARPS; adult respiratory distress syndrome; inflammatory bowel disease; Crohn's disease; ulcerative colitis; periodontal disease; granulocyte activation; chronic inflammation; allergy.

Homo sapiens.

W020022899-A2.

11-APR-2002.

03-OCT-2001; 2001WO-US030821.

03-OCT-2000; 2000US-0237189P.

(GENE-) GENE LOGIC INC.

Baezzer-Barclay Y, Weisman SM, Yamaga S, Vockley J; WPI; 2002-435328/46.

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity.

Claim 1; SEQ ID NO 1085; 114p; English.

The invention relates to detecting (M) granulocyte (GC) activation (GCA) by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARPS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/publicated_pct_sequences](http://wipo.int/pub/publicated_pct_sequences)

Sequence 12263 BP; 2497 A; 3568 C; 3664 G; 2534 T; 0 U; 0 Other;

13.6%; Score 362.4; DB 6; Length 12263;

Best Local Similarity		99.7%	Pred No. 3.6e-40;	
Matches	363;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
Qy	2187	GGGGGCCGTGCTGCCCTGCCCCAGAGCTGTGTCTTGATGCGCGTGTTCAGGCTCATGAGACAG	2246	
Db	11371	GGGGGCCGTCTGCCCTTCCAGAGCTGTGTCTTGATGCCGTGTTCAGGCTCATGAGACAG	1143	
Qy	2247	TGCTGGGGCTTATGAGAGCTGGGGCAGCGGCCCCAGCTTCAGACACACTTAAACAGAGACCTGCAG	2306	
Db	11431	TGCTGGGGCTTATGAGAGCTGGGGCAGCGGCCCCAGCTTCAGACACACTTAAACAGAGACCTGCAG	1149	
Qy	2307	AGCATCCGAAAGCGGACATCGGTAGAGGCTGGGACCCCTTCTCAAGCTGTGTGAGCTCTGCA	2366	
Db	11491	AGCATCCGAAAGCGGACATCGGTAGAGGCTGGGACCCCTTCTCAAGCTGTGTGAGCTCTGCA	1155	
Qy	2367	GGCTTAGTGCAGCTCTCTCAGCGAGCTTCAGCTCATATGCTAGACAGCTTTACAGTCTCTG	2426	
Db	11551	GGCTTAGTGCAGCTCTCTCAGCGAGCTTCAGCTCATATGCTAGACAGCTTTACAGTCTCTG	11611	
Qy	2427	GACTCTGCGCACAGCATTCACACTGCGGACAGATGACAGCGCGCTGTCTCTGTGTCT	2486	
Db	11611	GACTCTGCGCACAGCATTCACACTGCGGACAGATGACAGCGCGCTGTCTCTGTGTCT	1167	
Qy	2487	CTGCTGCTGCGCAGGGGTTCTCTTCCGGGACAGAAACAATAAACCACTGTGSCCACTG	2546	
Db	11671	CTGCTGCTGCGCAGGGGTTCTCTTCCGGGACAGAAACAATAAACCACTGTGSCCACTG	1173	
Qy	2547	AAAAA	2550	
Db	11731	AAACA	11734	
RESULT 12				
ID	ABA94501	standard; DNA; 15297 BP.		
XX	XX	ABA94501;		
AC	XX	09-APR-2002 (first entry)		
DT	XX			
DE	XX	Human proto-oncogene tyrosine kinase genomic DNA sequence.		
XX	XX			
XX	KW	Proto-oncogene tyrosine kinase; poTK; tumour; cytosolic; anti-leukemic;		
KW	KW	gene therapy; protein therapy; vaccine; enzymatic-inhibition; human;		
KW	KW	anti-kinase; ds.		
XX	OS	Homo sapiens.		
XX	XX			
XX	Key	Location/Qualifiers		
FT	CDS	2563..13081		
FT	FT	/*cag= a		
FT	FT	/product= "poTK"		
FT	FT	/note= "contains introns"		
FT	exon	2563..2275		
FT	FT	/*cag= b		
FT	FT	/number= 1		
FT	FT	2276..2927		
FT	FT	/*cag= c		
FT	FT	/number= 1		
FT	exon	2928..3101		
FT	FT	/*cag= d		
FT	FT	/number= 2		
FT	FT	3102..4476		
FT	FT	/*cag= e		
FT	FT	/number= 2		
FT	FT	4477..4573		
FT	FT	/*cag= f		
FT	FT	/number= 3		
FT	FT	4574..4702		
FT	FT	/*cag= g		
FT	FT	/number= 3		
FT	FT	4703..4886		
FT	FT	/*cag= h		

FT /number= 4
FT intron 4887..6827
FT /tag= 1
FT /number= 1
FT exon 6828..6965
FT /tag= j
FT /number= 5
FT intron 6966..7038
FT /tag= k
FT /number= 5
FT exon 7039..7158
FT /tag= 1
FT /number= 6
FT intron 7159..7361
FT /tag= m
FT /number= 6
FT exon 7362..7484
FT /tag= n
FT /number= 7
FT intron 7485..7614
FT /tag= o
FT /number= 7
FT exon 7615..7801
FT /tag= p
FT /number= 8
FT intron 7802..7923
FT /tag= q
FT /number= 8
FT exon 7924..8007
FT /tag= r
FT /number= 9
FT intron 8008..9075
FT /tag= s
FT /number= 9
FT exon 9076..9198
FT /tag= t
FT /number= 10
FT intron 9199..9579
FT /tag= u
FT /number= 10
FT exon 9580..9633
FT /tag= v
FT /number= 11
FT intron 9634..10229
FT /tag= w
FT /number= 11
FT exon 10230..10348
FT /tag= x
FT /number= 12
FT intron 10349..10622
FT /tag= y
FT /number= 12
FT exon 10623..10717
FT /tag= z
FT /number= 13
FT intron 10718..10813
FT /tag= aa
FT /number= 13
FT exon 10814..10937
FT /tag= ab
FT /number= 14
FT intron 10938..11176
FT /tag= ac
FT /number= 14
FT exon 11177..11334
FT /tag= ad
FT /number= 15
FT intron 11335..11458
FT /tag= ae
FT /number= 15
FT exon 11459..11581
FT /tag= af
FT /number= 16

FT intron 11582..12938
FT /tag= ag
FT /number= 16
FT exon 12939..13078
FT /tag= ah
FT /number= 17
XX US6340584-B1.
XX
XX 22-JAN-2002.
XX
XX 27-MAR-2001; 2001US-00817180.
XX
XX 27-MAR-2001; 2001US-00817180.
XX
XX (PEKE) PE CORP NY.
XX
XX Gan W, Ye J, Di Francesco V, Beasley EM,
XX WPI, 2002-138497/18.
XX P-PSDB; ABB07354.
XX
XX Nucleic acids encoding a proto-oncogene tyrosine kinase, useful for the
XX prevention, diagnosis and treatment of e.g. leukemia and lung tumors.
XX
XX Claim 1; Fig 3A-G; 49pp; English.
XX
XX The invention provides isolated nucleic acid sequences encoding a proto-
XX oncogene tyrosine kinase (pork). The pork polynucleotides and protein may
XX be used in the prevention, diagnosis and treatment of diseases associated
XX with inappropriate pork expression, such as lung and kidney tumors,
XX leukemia and stomach adenocarcinoma. pork may be used to treat disorders
XX associated with decreased expression by rectifying mutations or deletions
XX in a patient's genome that affect the activity of pork by expressing
XX inactive proteins or to supplement the patients own production of pork.
XX The encoded pork may be used as an antigen in the production of
XX antibodies against pork and in assays to identify modulators of pork
XX expression and activity. The anti-pork antibodies and antagonists may be
XX used to down regulate expression and activity and as diagnostic agents
XX for detecting the presence of pork in samples. The present sequence
XX represents the human pork genomic DNA sequence
XX
SQ Sequence 15297 BP; 3267 A; 4306 C; 4478 G; 3246 T; 0 U; 0 Other;
Query Match 13.6%; Score 362.4; DB 6; Length 15297;
Best Local Similarity 99.7%; Pred. No. 3.6e-40;
Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2187 GGGGGCCCTGCCCCGACAGCTGTGCTGATGCCGTGTCAGGCTCATGAGCAG 2246
DB 12938 GGGGGCCCTGCCCCGACAGCTGTGCTGATGCCGTGTCAGGCTCATGAGCAG 12997
QY 2247 TGCTGGGCTTATGAGCTGGGAGCGGCCAGCTTACAGCATTACAGAGCTGACG 2306
DB 12998 TGCTGGGCTTATGAGCTGGGAGCGGCCAGCTTACAGCATTACAGAGCTGACG 13057
QY 2307 AGCATCGAAGGGGAGTGAAGGCTGGAGACCCCTTCTCAAGCTGTGAGCTTGCAG 2366
DB 13058 AGCATCGAAGGGGAGTGAAGGCTGGAGACCCCTTCTCAAGCTGTGAGCTTGCAG 13117
QY 2367 GGCCTAAGTGCAGCTCTTCAAGGCGCTCCAGCTCATATGTGTGACAGCTTTCAAGTCCG 2426
DB 13118 GGCCTAAGTGCAGCTCTTCAAGGCGCTCCAGCTCATATGTGTGACAGCTTTCAAGTCCG 13177
QY 2427 GACTCTGCACAGCATCCACATGCGCGGAGATGACGCGCGTCTCTCTGTGTC 2486
DB 13178 GACTCTGCACAGCATCCACATGCGCGGAGATGACGCGCGTCTCTCTGTGTC 13237
QY 2487 CCTGTGCTGCCAGGCTTCTCTTCCGGGAGAAAACAATTAACCATTTGTGCCACTG 2546
DB 13238 CCTGTGCTGCCAGGCTTCTCTTCCGGGAGAAAACAATTAACCATTTGTGCCACTG 13297
QY 2547 AAAA 2550

FT /number= 15
 FT 11335..11458
 FT /tag= ap
 FT /number= 15
 FT exon 11459..11581
 FT /tag= ag
 FT /number= 16
 FT Intron 11582..12938
 FT /tag= at
 FT /number= 16
 FT exon 12939..13078
 FT /tag= as
 FT /number= 17
 XX MO200277191-A2.
 XX
 XX 03-OCT-2002.
 XX
 XX 27-MAR-2002; 2002MO-US009325.
 XX
 XX 27-MAR-2001; 2001US-00817180.
 XX 06-DEC-2001; 2001US-00003295.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Gan W, Ye J, Di Francesco V, Beasley EM,
 XX WPI; 2003-029927/02.
 XX P-PSDB; ABB99046.
 XX
 XX New human kinase protein, useful for treating or diagnosing disorders
 PT associated with an absence of, inappropriate, or unwanted expression of
 PT the protein, e.g. inflammation or cancer, in drug screening assays and
 PT pharmacogenomics.
 PT
 PT Claim 4 (b) ; Fig 3; 75bp; English.
 XX
 XX The invention relates to a newly isolated peptide sequence of a human
 CC kinase that is related to the proto-oncogene tyrosine kinase subfamily.
 CC The activity of the kinase of the invention may be described as,
 CC cytostatic, antiarteriosclerotic, antiinflammatory and antiproliferative.
 CC Peptides of the invention are useful in assays to determine the
 CC biological activity of the protein, in drug screening assays, tissue
 CC typing and pharmacogenomic analysis. They are also useful in treating or
 CC diagnosing disorders characterised by an absence of, inappropriate, or
 CC unwanted expression of the protein, such as inflammation, cancer (e.g.
 CC leukemia, lung tumours, kidney tumours or stomach adenocarcinoma),
 CC arteriosclerosis, and psoriasis. Nucleic acid intermediates of the invention
 CC are useful as probes, primers and chemical intermediates in biological
 CC assays. The peptide and nucleic acid sequences are useful as models for
 CC the development of human therapeutic targets, aid in the identification
 CC of therapeutic proteins and serve as targets for the development of human
 CC therapeutic agents that modulate kinase protein activity in cells and
 CC tissues that express the protein. The protein of the invention may also
 CC be useful in gene therapy. The gene encoding the protein of the invention
 CC has been localised to human chromosome 15. The current sequence
 CC represents the human kinase genomic DNA sequence
 CC
 CC
 SQ Sequence 15297 BP; 3267 A; 4306 C; 4478 G; 3246 T; 0 U; 0 Other;
 Query Match 13.6%; Score 362.4; DB 8; Length 15297;
 Best Local Similarity 99.7%; Pred. No. 3.6e-40;
 Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2187 GGGGGCGGTGCTGCTGCGCCAGAGCTGTCTGTATGCCGTGTTCAAGGCTCATGAGACG 2246
 DB 12938 GGGGGCGGTGCTGCTGCGCCAGAGCTGTCTGTATGCCGTGTTCAAGGCTCATGAGACG 12997
 QY 2247 TGGTGGGCTTATGAGCTTGAGCGGCGCCAGCTTGAGCACCATTACCGAGAGCTGCAG 2306
 DB 12998 TGGTGGGCTTATGAGCTTGAGCGGCGCCAGCTTGAGCACCATTACCGAGAGCTGCAG 13057
 QY 2307 AGCATCCGAAGGCGCATCGGTGAGGAGCCCTTCTCAAGCTGCGGCTCTGCA 2366

DB 13058 AGCATCCGAAGGCGCATCGGTGAGGAGCCCTTCTCAAGCTGCGGCTCTGCA 13117
 QY 2367 GGGCTTAGTGAGCTCTCTAGCGGGCTCCAGCTCATATGCTGACAGCTCTTACAGTCTTG 2426
 DB 13118 GGGCTTAGTGAGCTCTCTAGCGGGCTCCAGCTCATATGCTGACAGCTCTTACAGTCTTG 13177
 QY 2427 GAGTCTTGACACAGATACACATGCGCGGAGAGAGCGCGGTGCTCTGTCGTC 2486
 DB 13178 GAGTCTTGACACAGATACACATGCGCGGAGAGAGCGCGGTGCTCTGTCGTC 13237
 QY 2487 CCTGTGCTGCGAGGCTTCTCTTCCGGGAGAAACAATAAACAATTGTGTCACACTG 2546
 DB 13238 CCTGTGCTGCGAGGCTTCTCTTCCGGGAGAAACAATAAACAATTGTGTCACACTG 13297
 QY 2547 AAAA 2550
 DB 13298 AACA 13301
 RESULT 14
 ADM28582
 ID ADM28582 standard; cDNA; 15297 BP.
 XX
 XX AC ADM28582;
 XX
 XX 17-JUN-2004 (first entry)
 XX
 XX Human proto-oncogene tyrosine kinase DNA.
 XX
 XX de; gene; gene therapy; human; proto-oncogene; tyrosine kinase; cancer.
 XX
 XX Homo sapiens.
 XX
 XX
 FH Key
 FT variation
 FT /tag= a
 FT /note= "Single nucleotide polymorphism"
 FT replace(847,g)
 FT /tag= b
 FT /note= "Single nucleotide polymorphism"
 FT replace(2159,c)
 FT /tag= c
 FT /note= "Single nucleotide polymorphism"
 FT replace(2484,t)
 FT /tag= d
 FT /note= "Single nucleotide polymorphism"
 FT replace(2563,f)
 FT /tag= e
 FT /note= "Single nucleotide polymorphism"
 FT replace(2577,c)
 FT /tag= g
 FT /note= "Single nucleotide polymorphism"
 FT replace(2927,a)
 FT /tag= h
 FT /note= "Single nucleotide polymorphism"
 FT replace(2807..2809,gg)
 FT /tag= i
 FT /note= "Single nucleotide polymorphism"
 FT replace(2808,t)
 FT /tag= j
 FT /note= "Single nucleotide polymorphism"
 FT replace(2922,a)
 FT /tag= k
 FT /note= "Single nucleotide polymorphism"
 FT replace(2928..3101
 FT /tag= l
 FT /note= "Single nucleotide polymorphism"
 FT replace(3102..4476
 FT /tag= m
 FT /note= "Single nucleotide polymorphism"
 FT replace(4312,t)
 FT /tag= n
 FT /note= "Single nucleotide polymorphism"
 FT replace(4477..4573
 FT exon

FT intron /tag= o
 FT 4574..4702
 FT /tag= d
 FT 4703..4886
 FT /tag= q
 FT 4887..6827
 FT /tag= r
 FT replace(4903,a)
 FT /tag= s
 FT /note= "Single nucleotide polymorphism"
 FT replace(5193,g)
 FT /tag= t
 FT /note= "Single nucleotide polymorphism"
 FT 6828..6965
 FT /tag= u
 FT 6966..7038
 FT /tag= v
 FT 7039..7158
 FT /tag= w
 FT 7159..7361
 FT /tag= x
 FT 7362..7484
 FT /tag= y
 FT 7485..7614
 FT /tag= z
 FT 7615..7801
 FT /tag= aa
 FT 7802..7923
 FT /tag= ab
 FT 7924..8007
 FT /tag= ac
 FT 8008..9075
 FT /tag= ad
 FT 9076..9198
 FT /tag= ae
 FT 9199..9579
 FT /tag= af
 FT 9580..9633
 FT /tag= ag
 FT 9634..10229
 FT /tag= ah
 FT 10230..10348
 FT /tag= ai
 FT 10349..10622
 FT /tag= aj
 FT 10623..10717
 FT /tag= ak
 FT 10718..10813
 FT /tag= al
 FT 10814..10937
 FT /tag= am
 FT 10938..11176
 FT /tag= an
 FT 11177..11334
 FT /tag= ao
 FT 11335..11458
 FT /tag= ap
 FT 11459..11581
 FT /tag= aq
 FT 11582..12938
 FT /tag= ar
 FT 12939..13081
 FT /tag= as
 PN US2004063130-A1.
 PD 01-APR-2004.
 PF 12-SEP-2003; 2003US-00660763.
 PR 27-MAR-2001; 2001US-00817180.
 PR 06-DEC-2001; 2001US-00003295.
 XX

PA (APPL-) APPLERA CORP.
 XX
 PI Gan W, Ye J, Di Francesco V, Beasley EM;
 XX
 DR WPI, 2004-282461/26.
 DR P-PSDB; ADM28581.
 PT New human kinase peptides, useful for preparing a composition for
 PT treating a disease or condition mediated by human kinases.
 PT
 PS Claim 4; SEQ ID NO 3; 53pp; English.
 CC
 CC The invention relates to a new isolated human proto-oncogene tyrosine
 CC kinase peptide or its allelic variant, orthologue or fragment. The
 CC peptide is useful for preparing a composition for treating a disease or
 CC condition mediated by a human kinase protein e.g. cancer. The present
 CC sequence represents the human proto-oncogene tyrosine kinase DNA.
 XX
 SQ Sequence 15297 BP; 3267 A; 4306 C; 4478 G; 3246 T; 0 U; 0 Other;
 Query Match 13.6%; Score 362.4; DB 12; Length 15297;
 Best Local Similarity 99.7%; Pred. No. 3.6e-40;
 Matches 363; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2187 GGGGGCCGTCTGCCCCGACAGCTGTCTGATGCGGTTCAGGCTCATGAGCAG 2246
 DB 12938 GGGGGCCGTCTGCCCCGACAGCTGTCTGATGCGGTTCAGGCTCATGAGCAG 12997
 QY 2247 TGTCTGGGCTTATGACCTTGGGAGGGGCCCACTTCAGACATCTACAGAGCTGCAG 2306
 DB 12998 TGTCTGGGCTTATGACCTTGGGAGGGGCCCACTTCAGACATCTACAGAGCTGCAG 13057
 QY 2307 AGCATCCGAAAGCGGATCGGTGAGGAGGAGCCCTTCTCAAGCTGTGCTTCGA 2366
 DB 13058 AGCATCCGAAAGCGGATCGGTGAGGAGGAGCCCTTCTCAAGCTGTGCTTCGA 13117
 QY 2367 GGCCTAGGTGACGCTCTCAGCGGCTCCAGCTCATATGCTGACAGCTCTTCAAGTCTTG 2426
 DB 13118 GGCCTAGGTGACGCTCTCAGCGGCTCCAGCTCATATGCTGACAGCTCTTCAAGTCTTG 13177
 QY 2427 GACTCTGTCACACAGATTCACACTGCGGCGAGATGACGCCCGGTCTCTCTGTGTC 2486
 DB 13178 GACTCTGTCACACAGATTCACACTGCGGCGAGATGACGCCCGGTCTCTCTGTGTC 13237
 QY 2487 CCGTGTGTCGCAAGGCTTCTCTTCCGGGAGAAACATTAACCACTTGATGCCACTG 2546
 DB 13238 CCGTGTGTCGCAAGGCTTCTCTTCCGGGAGAAACATTAACCACTTGATGCCACTG 13297
 QY 2547 AAAA 2550
 DB 13298 AACA 13301
 RESULT 15
 ID ADL71054 standard; DNA; 2584 BP.
 XX ADL71054;
 AC
 XX 20-MAY-2004 (first entry)
 DT
 XX Gene encoding type II collagen expression promoting protein, seq id 47.
 DE
 XX Osteopathic; antiinflammatory; antirheumatic; antiarthritic;
 KW gene therapy; type II collagen; expression; cartilage disease;
 KW osteoarthritis; cartilage defect; rheumatoid arthritis; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003087375-A1.
 XX
 PD 23-OCT-2003.
 XX

PF 16-APR-2003; 2003MO-JP004802.
 XX 16-APR-2002; 2002JP-0013908.
 PR 19-APR-2002; 2002US-0373594P.
 XX (ASAH) ASAH KASEI KK.
 PA Matsuda A, Honda G, Muramatsu S;
 PI MPI; 2003-045331/78.
 DR P-SDB; ADL71055.
 XX
 XX New purified protein that promotes type II collagen expression, useful
 PT for preventing and treating a cartilage disease, e.g. osteoarthritis,
 PT cartilage defect, or rheumatoid arthritis.
 XX
 XX Claim 4; SEQ ID NO 47; 271bp; English.
 CC The invention relates to a purified protein (I) that promotes type II
 CC collagen expression. Also disclosed is an isolated polynucleotide
 CC encoding (I), a recombinant vector comprising the polynucleotide and a
 CC gene therapeutic agent comprising the recombinant vector as an active
 CC ingredient. The proteins, genes, agents and methods are useful for
 CC preventing and treating a cartilage disease, e.g. osteoarthritis,
 CC cartilage defect, or rheumatoid arthritis. The current sequence
 CC represents a human gene of the invention encoding a protein that promotes
 CC type II collagen expression.
 CC
 SQ Sequence 2584 BP; 922 A; 446 C; 565 G; 651 T; 0 U; 0 Other;
 Query Match 12.3%; Score 328.8; DB 11; Length 2584;
 Best Local Similarity 62.0%; Pred. No. 1.3e-35;
 Matches 573; Conservative 0; Mismatches 342; Indels 9; Gaps 3;
 QY 1392 AACCTGTACCGATGGAAGGAGGCTTCTAGCATCTTGTGCTCATGACCACTTA 1451
 DB 1552 AACATGTATGATTCGAGGCGACTGGGTTTCAAACTTCTCACTTAATGATCATAC 1611
 QY 1452 CTGAGCACCAGACGCCCTCACCAGAGAGTGTGTCTCTGCAAGGCGTGTGCC 1511
 DB 1612 TATACAAACAAACAGGTCATCAAGAAATCAGTGTAGTCTGCTGAATCCTATTCCT 1671
 QY 1512 AAGGACAG---TGGGCTGGAACCATGAGACCTGTGTGGTGAAGCATTTGACAGG 1568
 DB 1672 AAGGACAGAAATGATTTCTAGTCATGAAATGTCATATTTGAGAAATTAATCTGGGCAAG 1731
 QY 1569 GGGAACTTTGGCGAAGTGTTCAGCGAGCGCTGCGAGCCGACAAACCTGTGGCGGTG 1628
 DB 1732 GGAATTTTGTGTAAGTATTAAG---GCACATTAAGATTAATACTTGTTCCTGTT 1788
 QY 1629 AAGCTTTGTGAGAGACGCTCCCACTGACCTCAAGGCCAAGTTTCTACAGAAAGCGAGG 1688
 DB 1789 AAAACATGTAAAGATTTCTCTCAGGAATGAAATTAATAATTTTACAAAGACCAAA 1848
 QY 1689 ATCTGAAGAGTACAGCCACCCCAATGTGTGCTCATTTGTGTCTGACCCAGAG 1748
 DB 1849 ATTCTCAAGCAATATATATATCCCAATATGTCAAACTTATAGAGTTTGACACAAAGA 1908
 QY 1749 CAGCCCATCTACATGTCATGAGCTTGTGCAAGGGGGGCGACTTCTGACCTTCTCCGC 1808
 DB 1909 CAGCCTGTCTACATCATTAATGAGTGTGTCAGAGGTATTTCTCACCCTTTCTGAGA 1968
 QY 1809 ACGGAGGGGCGCGCTGCGGGTGAAGACTGTGTCAGATGTGGGGATGCAAGCTGCT 1868
 DB 1969 AGGAGAGAGATGAATTAATACTCAACAGTATGTAATTTTCAATTAAGCGCTGCTGCT 2028
 QY 1869 GGCATGAGATACCTGAGAGCAAGTGTGATCCACCGGACCTGTGGTGTCCGAACTGC 1928
 DB 2029 GGTATGTTGTATCTCGAGATTAATACTGTATACAGGGACCTTGTGCAAGAACTGC 2088
 QY 1929 CTGGTGA CAGAGAAATGTCTGTAAGATCAGTGACTTTGGAGTGTCCGAGAGGAGCC 1988
 DB 2089 CTGGTAGGTAAATAATTTCTGAAAATCAGTGAATTTGGAAATGTCTGCTCAAGAGAT 2148

QY 1989 GATGGGCTATATGACACCTTCAGGGGGCTTCAGACAAAGTCCCGTGAATGAGACCGCACTT 2048
 DB 2149 GGTGAGGTATATCATCTTC---TGGCTTAAGCAAGATTCCTTAATGAGACGACCG 2205
 QY 2049 GAGGCCCTTAACATGACGGCGCTACTCTCCGAAAGGACGTGTGAGCTTTGGCATCTTG 2108
 DB 2206 GAGGCTTATATATATGAGATACATTAAGAGATGAGTGTGAGCTTTGGCATCTT 2265
 QY 2109 CTCTGGAGACCTTCAGCCTTGGGGGCTCCCTATCCAACTTCAGCAATCAGACACA 2168
 DB 2266 CTCTGGAGACCTTCAGCCTTGAAGGCTTGTGCGTACCTGGAATGACAAATCAGACACA 2325
 QY 2169 CGGAGTTTGTGAGAGAGGGGGCGCTGTGCTGCGCCAGAGCTGTGTCTGATGCCGTG 2228
 DB 2326 AGAGACAAAGTAAAGAGATTAACCGAATGTCTGCTCCCAAGCATGTCCAGAGATAT 2385
 QY 2229 TTCAAGCTATGAGAGAGTGTGGGCTATGAGCTTGGGAGCGGCCAGCTTCAGCAACC 2288
 DB 2386 TCCAAATATCATGATGAAAGTGTGGGATTATTAACCTGAAAATGCGCTTAATTCAGTGAA 2445
 QY 2289 ATCTACAGAGCTGCAAGCATC 2312
 DB 2446 CTTCAGAAAGAGCTCATATCATC 2469

Search completed: March 19, 2005, 04:31:15
 Job time : 1290.71 secs

This Page Blank (uspto)

Db 241 GGGGCCAGACCGGGCCATCAGCCCTGACAGCCCATCAGTCACTGGGCTGAGATCA 300
QY 301 CCAGCCAACTAGAGGCGCTGAGCCGCTTGTCTGGGAGACAGCAGAGATCTGAATCAG 360
Db 301 CCAGCCAACTAGAGGCGCTGAGCCGCTTGTCTGGGAGACAGCAGAGATCTGMACTCAG 360
QY 361 GGGCCCTGAGCAGTGAAGCTGAGTCACTCGGGGAAAGGAGCAGCTTTCGAAAGCTTACA 420
Db 361 GGGCCCTGAGCAGTGAAGCTGAGTCACTCGGGGAAAGGAGCAGCTTTCGAAAGCTTACA 420
QY 421 GCGAGCAGTGAAGCAGCAGCTGACAGAGAGTCACTCAAGACCAAGCAGCAGCATTGAGA 480
Db 421 GCGAGCAGTGAAGCAGCAGCTGACAGAGAGTCACTCAAGACCAAGCAGCAGCATTGAGA 480
QY 481 AGCTGAAGACCGTACCGAGCTCTGGCACGGGACAGTCCCAAGCCAAAGCCAAAGTACC 540
Db 481 AGCTGAAGACCGTACCGAGCTCTGGCACGGGACAGTCCCAAGCCAAAGCCAAAGTACC 540
QY 541 AGGAGGCGAGCAAGAACAAGGACCGTGAACAAGGCAAGATATGTGCGCAGCTGT 600
Db 541 AGGAGGCGAGCAAGAACAAGGACCGTGAACAAGGCAAGATATGTGCGCAGCTGT 600
QY 601 GGAAGCTCTTGTCTCAACAACCGCTATGTGTGGCGGTGCGGAGCTGCGACACTACAC 660
Db 601 GGAAGCTCTTGTCTCAACAACCGCTATGTGTGGCGGTGCGGAGCTGCGACACTACAC 660
QY 661 ACCAGCACCAACCAAGCTCTGCTGCGCGCTGCTGCGGTCACTGACAGACTTGCAG 720
Db 661 ACCAGCACCAACCAAGCTCTGCTGCGCGCTGCTGCGGTCACTGACAGACTTGCAG 720
QY 721 AGGAGATGAGCTTCACTCTGAAGAAGATCTGAGAGAAATACCTGAGAGATTAGAGCTGG 780
Db 721 AGGAGATGAGCTTCACTCTGAAGAAGATCTGAGAGAAATACCTGAGAGATTAGAGCTGG 780
QY 781 TGCAGATGAGGTGTGTGGCCATTCAACGGAGATGAGCTGCTGCCCGATCAAG 840
Db 781 TGCAGATGAGGTGTGTGGCCATTCAACGGAGATGAGCTGCTGCCCGATCAAG 840
QY 841 CTGAGGCTGAGTACCAAGGCTTCTGAGAGAGATGTGGTCCGACCTGAGAGTCCCACT 900
Db 841 CTGAGGCTGAGTACCAAGGCTTCTGAGAGAGATGTGGTCCGACCTGAGAGTCCCACT 900
QY 901 GTGTCAAGTTCAGATGAGTCACTGCTTGAAGAGAGGTGAACCGCTGAGAGCTGGGAGCTCC 960
Db 901 GTGTCAAGTTCAGATGAGTCACTGCTTGAAGAGAGGTGAACCGCTGAGAGCTGGGAGCTCC 960
QY 961 AGCTGAACGAGCTGACTGTGAGAGAGCGTGCAGACAAGCTGACCTCAGTGAACAGATGAGC 1020
Db 961 AGCTGAACGAGCTGACTGTGAGAGAGCGTGCAGACAAGCTGACCTCAGTGAACAGATGAGC 1020
QY 1021 TGGCTGTGTGACACCGAGATGTGTGTTACAGGCGGAGAGAGATGTGTTAGCAGCTGCAACAGG 1080
Db 1021 TGGCTGTGTGACACCGAGATGTGTGTTACAGGCGGAGAGAGATGTGTTAGCAGCTGCAACAGG 1080
QY 1081 AGCTCCGGAATGAAGAGAGAAACCAACCCCGGAGCGGGGTGACAGCTGTGGCAAGA 1140
Db 1081 AGCTCCGGAATGAAGAGAGAAACCAACCCCGGAGCGGGGTGACAGCTGTGGCAAGA 1140
QY 1141 GGGCAAGTGTGCAAGAAGCACTGCAAGGGGTGTGCAAGCTGTGCAAGCAGGCAAGC 1200
Db 1141 GGGCAAGTGTGCAAGAAGCACTGCAAGGGGTGTGCAAGCTGTGCAAGCAGGCAAGC 1200
QY 1201 TGCAGGCGCCAGCAGAGATTGCTGACAGCAAGCTGAGAGCACTGGGCGCCGAGGAGCCCC 1260
Db 1201 TGCAGGCGCCAGCAGAGATTGCTGACAGCAAGCTGAGAGCACTGGGCGCCGAGGAGCCCC 1260
QY 1261 CGGCTGTGTGCTCTGTCAGAGATGACCGCACTTCAGGTGTGTCTGAGAGAGAGAGAG 1320
Db 1261 CGGCTGTGTGCTCTGTCAGAGATGACCGCACTTCAGGTGTGTCTGAGAGAGAGAGAG 1320
QY 1321 AGGGGGAGAGAGACCAACCGCTGAGATCTTAAAGCCCAATCTCAGGAATCTTCGCGC 1380
Db 1321 AGGGGGAGAGAGACCAACCGCTGAGATCTTAAAGCCCAATCTCAGGAATCTTCGCGC 1380

Db 1321 AGGGGGAGAGAGACCAACCGCTGAGATCTTAAAGCCCAATCTCAGGAATCTTCGCGC 1380
QY 1381 CCAAGTTCCTGAACCTGTATCCGATCTGGAAGGGGAGGCTTTCCTGACATTCCTTGTGCTA 1440
Db 1381 CCAAGTTCCTGAACCTGTATCCGATCTGGAAGGGGAGGCTTTCCTGACATTCCTTGTGCTA 1440
QY 1441 TGCACCACTACTGAGCAACCCAGAGCCCTCACAAGAAAGATGTGTGTCCTGACAGA 1500
Db 1441 TGCACCACTACTGAGCAACCCAGAGCCCTCACAAGAAAGATGTGTGTCCTGACAGA 1500
QY 1501 GGGCTGTGTCCAAAGACAAAGTGGTGTGTAACATAGAGACCTGTGTGGTGAAGAGA 1560
Db 1501 GGGCTGTGTCCAAAGACAAAGTGGTGTGTAACATAGAGACCTGTGTGGTGAAGAGA 1560
QY 1561 TTGAGACGGGGGAACTTTGGCGAAGTGTTCAGGGAGCGCTGGAGCGCAACACCTGG 1620
Db 1561 TTGAGACGGGGGAACTTTGGCGAAGTGTTCAGGGAGCGCTGGAGCGCAACACCTGG 1620
QY 1621 TGGCGGTGAAGTCTTGTGAGAGAGCGTCCACCTGACCTCAAGGCAAGTTTCTACAG 1680
Db 1621 TGGCGGTGAAGTCTTGTGAGAGAGCGTCCACCTGACCTCAAGGCAAGTTTCTACAG 1680
QY 1681 AAGCAGAGATCTTGAAGCAGTACAGCCACCAACATGCTGTCTCATTTGTGTCTGA 1740
Db 1681 AAGCAGAGATCTTGAAGCAGTACAGCCACCAACATGCTGTCTCATTTGTGTCTGA 1740
QY 1741 CCAGAGAGAGCCCAATCTATATGTATAGAGCTTGTGAGGGGGGCGACTTCTGACCT 1800
Db 1741 CCAGAGAGAGCCCAATCTATATGTATAGAGCTTGTGAGGGGGGCGACTTCTGACCT 1800
QY 1801 TCTCTCGCAAGAGGGGGGCGCGCTGCGGTGAAGACTGTGTGACAGATGTGTGGGAGATG 1860
Db 1801 TCTCTCGCAAGAGGGGGGCGCGCTGCGGTGAAGACTGTGTGACAGATGTGTGGGAGATG 1860
QY 1861 CAGCTGTGTGCAATGAGATCTCTGAGAGAGCAATGTCTGACACCGGAGCTGGGCTGCTC 1920
Db 1861 CAGCTGTGTGCAATGAGATCTCTGAGAGAGCAATGTCTGACACCGGAGCTGGGCTGCTC 1920
QY 1921 GGAAGCTGCTGTGACAGAGAAATGTCTGAAGATCAGTGACTTTGGGAGTGTCCGAG 1980
Db 1921 GGAAGCTGCTGTGACAGAGAAATGTCTGAAGATCAGTGACTTTGGGAGTGTCCGAG 1980
QY 1981 AGGAAGCCGATGGGGTCTATGAGAGCTCAAGGGGCGCTCAGACAAAGTCCCGTGAAGTGA 2040
Db 1981 AGGAAGCCGATGGGGTCTATGAGAGCTCAAGGGGCGCTCAGACAAAGTCCCGTGAAGTGA 2040
QY 2041 CCGCACCTGAGGCGCTTAACTACGGCGCTACTCTCTCGAAAGCAGCAGTGTGAGCTTGG 2100
Db 2041 CCGCACCTGAGGCGCTTAACTACGGCGCTACTCTCTCGAAAGCAGCAGTGTGAGCTTGG 2100
QY 2101 GGAATGTGTCTGGGAGACCTTCAAGCTGGGGGCGCTCCGCTATCCCAACCTTCAGCAATC 2160
Db 2101 GGAATGTGTCTGGGAGACCTTCAAGCTGGGGGCGCTCCGCTATCCCAACCTTCAGCAATC 2160
QY 2161 AGCAGACACGGGAGATTGTGAGAAAGGGGGCGGTCTGCTGCTGCCAGAGCTGTCTTG 2220
Db 2161 AGCAGACACGGGAGATTGTGAGAAAGGGGGCGGTCTGCTGCTGCCAGAGCTGTCTTG 2220
QY 2221 ATGCGGTGTTCAGGCTCATGAGACAGTGTGGGCTTATAGAGCTGGGAGCGGCGCAAGT 2280
Db 2221 ATGCGGTGTTCAGGCTCATGAGACAGTGTGGGCTTATAGAGCTGGGAGCGGCGCAAGT 2280
QY 2281 ATGCGGTGTTCAGGCTCATGAGACAGTGTGGGCTTATAGAGCTGGGAGCGGCGCAAGT 2280
Db 2281 ATGCGGTGTTCAGGCTCATGAGACAGTGTGGGCTTATAGAGCTGGGAGCGGCGCAAGT 2280
QY 2281 TGAAGACCATCTACAGAGAGCTGACAGAGATCGAAAGGGGAGTGGTGAAGCTGGGAGC 2340
Db 2281 TGAAGACCATCTACAGAGAGCTGACAGAGATCGAAAGGGGAGTGGTGAAGCTGGGAGC 2340
QY 2341 CCCTTCTCAAGCTGTGTGCTCTGAGAGCTTGAAGTGAAGTCTCTCAAGGCTTCAAGTCA 2400
Db 2341 CCCTTCTCAAGCTGTGTGCTCTGAGAGCTTGAAGTGAAGTCTCTCAAGGCTTCAAGTCA 2400
QY 2401 TATGTGACAGCTCTTACAGATCTGAGACTCTGCTCCACAGCATTCACACTGCGGAGAG 2460
Db 2401 TATGTGACAGCTCTTACAGATCTGAGACTCTGCTCCACAGCATTCACACTGCGGAGAG 2460

QY 2461 ATGACAGGCGGTGCTCTGTGTGTCCTGCTGCGAGGCTTCTCTTCCGGGCAAG 2520
DB 2461 ATGACAGGCGGTGCTCTGTGTGTCCTGCTGCGAGGCTTCTCTTCCGGGCAAG 2520
QY 2521 AAGCAATAAACCACTTGTGCGCACTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2580
DB 2521 AAGCAATAAACCACTTGTGCGCACTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2580
QY 2581 AA 2640
DB 2581 AA 2640
QY 2641 AA 2674
DB 2641 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2674

RESULT 2
AR456323 2674 bp DNA linear PAT 20-FEB-2004
LOCUS AR456323
DEFINITION Sequence 1 from patent US 6686187.
ACCESSION AR456323
VERSION AR456323.1 GI:42691378
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2674)
AUTHORS Gan W., Ye J., Di Francesco V. and Beasley E.M.
TITLE Isolated human kinase proteins, and uses thereof
JOURNAL human kinase proteins, and uses thereof
FEATURES
source Location/Qualifiers
1..2674
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 2674; DB 6; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2674; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCGGGGTCCGCAACCGGGCTTGAAGTCCGTCGAGGCTTCCGAGACAGTCCCGTGC 60
DB 1 TCCGGGGTCCGCAACCGGGCTTGAAGTCCGTCGAGGCTTCCGAGACAGTCCCGTGC 60
QY 61 GGAACAGCACTATGGGCTTCTCTTGAAGTGAAGGCGGCGGCGGCGGCGGCGGCGG 120
DB 61 GGAACAGCACTATGGGCTTCTCTTGAAGTGAAGGCGGCGGCGGCGGCGGCGGCGG 120
QY 121 AGCAATGACAGAGGCGGAGCTTCTCTTGAAGTGAAGGCGGCGGCGGCGGCGGCGG 180
DB 121 AGCAATGACAGAGGCGGAGCTTCTCTTGAAGTGAAGGCGGCGGCGGCGGCGGCGG 180
QY 181 GGGTCAAGATGACAGAGGAGTATGACAGAGTCTTCAACAATGTCCTGCAAGACATG 240
DB 181 GGGTCAAGATGACAGAGGAGTATGACAGAGTCTTCAACAATGTCCTGCAAGACATG 240
QY 241 GGGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 241 GGGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 301 CAGAGCCAACTGAGGCGGCTTGAAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
DB 301 CAGAGCCAACTGAGGCGGCTTGAAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 361 GGGCCCTGAGGAGTGAAGCTTGTCTATCCGGGAAAGGCGGCGGCGGCGGCGGCGG 420
DB 361 GGGCCCTGAGGAGTGAAGCTTGTCTATCCGGGAAAGGCGGCGGCGGCGGCGGCGG 420
QY 421 GCGAGAGTGGGCGGAGCTGACAGAGGCTCAACAAGCCCAAGCGGAGGAGCTTGAAGA 480
DB 421 GCGAGAGTGGGCGGAGCTGACAGAGGCTCAACAAGCCCAAGCGGAGGAGCTTGAAGA 480

DB 421 GCGAGAGTGGGCGGAGCTGACAGAGGCTCAACAAGCCCAAGCGGAGGAGCTTGAAGA 480
QY 481 AGCTGAAGAGCCAGTACCGAGCTTGGGCAAGTGGCCCAAGGCAAGGCAATATAC 540
DB 481 AGCTGAAGAGCCAGTACCGAGCTTGGGCAAGTGGCCCAAGGCAAGGCAATATAC 540
QY 541 AAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 541 AAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 601 GGAAGCTCTTGTGTCACCAACCGCTATGTGTGCGGCGGCGGCGGCGGCGGCGGCGG 660
DB 601 GGAAGCTCTTGTGTCACCAACCGCTATGTGTGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 661 ACCAGACCAACCAACGCTCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 661 ACCAGACCAACCAACGCTCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 721 AAGAGATGGCTTGCATCTGAAAGAGATCTGAGGAGATTAAGGATTAAGGAGCTGG 780
DB 721 AAGAGATGGCTTGCATCTGAAAGAGATCTGAGGAGATTAAGGAGATTAAGGAGCTGG 780
QY 781 TGCAGAGTGAAGTGTGTCATTCACCGGAGATGAGTCAAGCTGCTGCGGCGGCGGCGG 840
DB 781 TGCAGAGTGAAGTGTGTCATTCACCGGAGATGAGTCAAGCTGCTGCGGCGGCGGCGG 840
QY 841 CTGAGGCTGATGACCAAGGCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 841 CTGAGGCTGATGACCAAGGCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
QY 901 GGTGCAAGTTCGATGAGTCACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
DB 901 GGTGCAAGTTCGATGAGTCACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
QY 961 AGCTGAACAGAGCTGATCTGAGAGGCGTCAAGCAACGCTGACCTGAGTACAGATGAG 1020
DB 961 AGCTGAACAGAGCTGATCTGAGAGGCGTCAAGCAACGCTGACCTGAGTACAGATGAG 1020
QY 1021 TGGCTGTGCGCACCGAGATGATGTTCAAGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 1021 TGGCTGTGCGCACCGAGATGATGTTCAAGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
QY 1081 AGCTCCGAAATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
DB 1081 AGCTCCGAAATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
QY 1141 GGCAGTGTGCAAGAGGAGCTGCAAGGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1200
DB 1141 GGCAGTGTGCAAGAGGAGCTGCAAGGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1200
QY 1201 TGCAGGCGGAGGAGGAGTGTGTCAGACCAAGCTGAGGAGGAGGAGGAGGAGGAG 1260
DB 1201 TGCAGGCGGAGGAGGAGTGTGTCAGACCAAGCTGAGGAGGAGGAGGAGGAGGAG 1260
QY 1261 CGCTGTGTGCTGCTGCAAGATGACCGGCACTTCAAGTGTGCTGCAAGGAGGAGGAG 1320
DB 1261 CGCTGTGTGCTGCTGCAAGATGACCGGCACTTCAAGTGTGCTGCAAGGAGGAGGAG 1320
QY 1321 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
DB 1321 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
QY 1381 CCAAGTCTTGAACCTGTAACCGAGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
DB 1381 CCAAGTCTTGAACCTGTAACCGAGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 TCGACCACTTACGAGACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1441 TCGACCACTTACGAGACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 1501 GGGCTGTGCGCAAGAGCAAGTGGGTCTGAACCATGAGGAGCTGTGTGTGAGTGAAGAG 1560
DB 1501 GGGCTGTGCGCAAGAGCAAGTGGGTCTGAACCATGAGGAGCTGTGTGTGAGTGAAGAG 1560


```

QY 1412 GGAAGGCTTTCCTAGCATTCCTTGTGCTCATTCAGACCACTTACTGAGGACCAGAGCCCTT 1471
DB 1623 GGAAGGCTTTCCTAGCATTCCTTGTGCTCATTCAGACCACTTACTGAGGACCAGAGCCCTT 1682
QY 1472 CACCAAGAGAAGTGTGTCTCTGTCACAGAGGCTGTGCCAAGAGCAAGTGGTGTGAA 1531
DB 1683 CACCAAGAGAAGTGTGTCTCTGTCACAGAGGCTGTGCCAAGAGCAAGTGGTGTGAA 1742
QY 1532 CCATGAGACCTGTGTGTGGTGTGAGCAGATTTGACCGGGGGAATTTGGCGAAGTGTTCAG 1591
DB 1743 CCATGAGACCTGTGTGTGGTGTGAGCAGATTTGACCGGGGGAATTTGGCGAAGTGTTCAG 1802
QY 1592 CGGAGCCCTGCGAGCCGACCAACACCCCTGTGTGAGTGTGTGAGAGAGCGCTCC 1651
DB 1803 CGGAGCCCTGCGAGCCGACCAACACCCCTGTGTGAGTGTGTGAGAGAGCGCTCC 1862
QY 1652 ACCTGACCTTCAAGGCGCAAGTTTCTACAGGAGCGAGATCTCTGAGCAGTACAGCCACC 1711
DB 1863 ACCTGACCTTCAAGGCGCAAGTTTCTACAGGAGCGAGATCTCTGAGCAGTACAGCCACC 1922
QY 1712 CAACATCGTGTCTCTCATTTGTGTCTGTCACCCAGAAAGCAGCCATCTACATGCTCATGGA 1771
DB 1923 CAACATCGTGTCTCTCATTTGTGTCTGTCACCCAGAAAGCAGCCATCTACATGCTCATGGA 1982
QY 1772 GCTTGTGACAGGGGGGCGACTTCTCTGACCTTCTCCGCAAGGAGGGGGCCGCTGCGGGT 1831
DB 1983 GCTTGTGACAGGGGGGCGACTTCTCTGACCTTCTCCGCAAGGAGGGGGCCGCTGCGGGT 2042
QY 1832 GAAGACTCTGCTGTCAGATGTGTGGGGAGATGACGCTGTGCTGAGTACCTGAGAGCAA 1891
DB 2043 GAAGACTCTGCTGTCAGATGTGTGGGGAGATGACGCTGTGCTGAGTACCTGAGAGCAA 2102
QY 1892 GTGCTGTGATCCACGGGAGCTGTGCTGTGCGGAACTGCTGTGTGACAGAGAGAAATGCTCT 1951
DB 2103 GTGCTGTGATCCACGGGAGCTGTGCTGTGCGGAACTGCTGTGTGACAGAGAGAAATGCTCT 2162
QY 1952 GAAGACTCTGATCTTGTGGAGTGTCCGAGAGAGAGCCGATGAGGCTTATGACAGCTCAGG 2011
DB 2163 GAAGACTCTGATCTTGTGGAGTGTCCGAGAGAGAGCCGATGAGGCTTATGACAGCTCAGG 2222
QY 2012 GGGCTCTCAGACCAAGTCCCCGCTGAGTGTGACCGGACCTGAGGCCCTTAACTACGCGCGCTA 2071
DB 2223 GGGCTCTCAGACCAAGTCCCCGCTGAGTGTGACCGGACCTGAGGCCCTTAACTACGCGCGCTA 2282
QY 2072 CTCTCTCCGAAAGGAGCTGTGTGAGCTTTGGCATCTTGTCTTGGGAGACCTTACGCTGGG 2131
DB 2283 CTCTCTCCGAAAGGAGCTGTGTGAGCTTTGGCATCTTGTCTTGGGAGACCTTACGCTGGG 2342
QY 2132 GGGCTCCCCCTTATCCCACTCAGACATCAGAGACACGGGAGTTTGTGGAGAAAGGGGG 2191
DB 2343 GGGCTCCCCCTTATCCCACTCAGACATCAGAGACACGGGAGTTTGTGGAGAAAGGGGG 2402
QY 2192 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2251
DB 2403 CCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2462
QY 2252 GGGCTTATGAGCTGTGGGAGGAGGCGCCAGCTTACAGACCATTTACAGAGAGTGTGAGAGCAT 2311
DB 2463 GGGCTTATGAGCTGTGGGAGGAGGCGCCAGCTTACAGACCATTTACAGAGAGTGTGAGAGCAT 2522
QY 2312 CCGAAGAGGAGCATCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2371
DB 2523 CCGAAGAGGAGCATCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2582
QY 2372 AGGTGACGCTCTCAGAGGCTCAGCTCATATGCTGACAGCTTTCACAGTCTTGAAGCTC 2431
DB 2583 AGGTGACGCTCTCAGAGGCTCAGCTCATATGCTGACAGCTTTCACAGTCTTGAAGCTC 2492
QY 2432 CTGCGACACAGATCCACATGCGCGGAGAGATGAGAGGCGCTGTCTCTGTGTCTCTGTGCT 2491
DB 2643 CTGCGACACAGATCCACATGCGCGGAGAGATGAGAGGCGCGCTGTGTCTCTGTGTCTCTGTGCT 2702

```

```

QY 2492 TGTGTCAGAGGCTTCTCTCTCCGAGCAGAAACATTAACCACTTGTGCCACTGAAAAA 2551
DB 2703 TGTGTCAGAGGCTTCTCTCTCCGAGCAGAAACATTAACCACTTGTGCCACTGAAAAA 2762
QY 2552 AAAAAAAAAAAAAAAAAAAAAA 2577
DB 2763 AAAAAAAAAAAAAAAAAAAAAA 2788

RESULT 4
HSCFES
LOCUS HSCFES 2771 bp mRNA linear PRI 01-DEC-1992
DEFINITION H.sapiens RNA for c-fes.
ACCESSION X52192
VERSION X52192.1 GI:29890
KEYWORDS fes cellular oncogene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 2771)
Alcalay, M., Antolini, F., Van de Ven, W.J., Lanfrancone, L.,
Grisman, F., and Pelicci, P.G.
Characterization of human and mouse c-fes cDNA clones and
identification of the 5' end of the gene
Oncogene 5 (3), 267-275 (1990)
JOURNAL 90191711
MEDLINE
PUBMED 2179816

FEATURES
source
location/Qualifiers
1..2771
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2771
/gene="c-fes"
76..2544
/gene="c-fes"
/codon_start=1
/protein_id="CAA36438.1"
/db_xref="GI:29891"
/db_xref="GDB:P07332"
/db_xref="UniProt/Swiss-Prot:P07332"
/translatability="MGFSESLCSPOGHVLOQMOBAELRLBGRKMAQRYKSDREY
AGLHMSLSDSGQSPRAISPDSPISQSAEITSGEISRLRQAEEDNLSPLSKL
SLIREROOLRKYSEMOOQOELVKTSHSOIEKLKSQYRALRASAQKRYQAS
KDXDRDAXKDXRYSRLIMKLFAHNNYVIVGVRALQHLHHOHLPLGLRSLODHEE
NACTIKETLOETLEISLVQDEYVAIHEMMAAARIQREBAIYQGLRQYGAAPVP
CVTFESLIEBEPLEPELOLNELVESVOHTLSVDELAVATEWFRKQEMVQL
QOELNBEENBHPERVOLGKROVLOELQVALCSQALQAOELLOQLTLEHIG
PEEPVLLLODRISTSSBOEREGGRTPLLEIKSHISGIFRPFSLPPLQILPE
VOKPLHEOLMTHGAIPLAEVABLVSAGDFLVRESQKQEVYLSVMDGLPHFIQS
LDNLVRLBEGEPSPILIDHLSVQOPLTKSGVILHRAVPRKQVNLNHEDLVGEQ
IGRGNFGEYSRRLADNTLVAVKSCRETLPLDLAKPLQEARILKQYSHNPVRLIG
VCTOKPLVITVELVQGGDFLFLRTBEGARLVKTLQVNGDAAGMEVLSCCILHR
DLAANCLVTEKNVILKSIDPFMSREBADDVVAASGSRQVPAKMPAPEALNYSRGSSE
SLVMEFGILMETFSLGASPYNLSNQOTREFVEKGRILPCBELCPDAVRLMEQCA
YRPGORPSFSPTIYELQSLRKHR"
2738..2743
/gene="c-fes"

ORIGIN
Query Match 87.1%; Score 2328; DB 9; Length 2771;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 2551; Conservative 0; Mismatches 5; Indels 210; Gaps 1;

QY 2 CCGGGGTCGCAACCGGGCTGAGTGTGCTGAGAGCCGTCCAGAGACGACTGCGCTGG 61
DB 6 CCGGGGTCGCAACCGGGCTGAGTGTGCTGAGAGCCGTCCAGAGACGACTGCGCTGG 65
QY 62 GAACAGACTATAGGCTTCTCTTGTGAGTGTGAGAGCCCGGAGGCGCAGGGGTCCTGCA 121
DB 66 GAACAGACTATAGGCTTCTCTTGTGAGTGTGAGAGCCCGGAGGCGCAGGGGTCCTGCA 125

```

122 GCAATGTCAGAGGCGGAGCTTGCTGTACTGAGGGGCA TGAGAAAGTGGATGGCCGAGG 181
126 GCAATGTCAGAGGCGGAGCTTGCTGTACTGAGGGGCA TGAGAAAGTGGATGGCCGAGG 185
182 GGTCAAGATGACAGGAGATATGACAGACTGCTTCA CCAATGTCCCTGACAGACAGTGG 241
186 GGTCAAGATGACAGGAGATATGACAGACTGCTTCA CCAATGTCCCTGACAGACAGTGG 245
242 GGGCCGAGGCGGAGCCATCAAGCCCTGACAGCCCATCA GTCACTCTGGGCTGAGATCAC 301
246 GGGCCGAGGCGGAGCCATCAAGCCCTGACAGCCCATCA GTCACTCTGGGCTGAGATCAC 305
302 CAGCCAACTGAGGGGCTGAGCGCTTGCTGAGGGGCA GCA CGAGAGATCTGAATCAGG 361
306 CAGCCAACTGAGGGGCTGAGCGCTTGCTGAGGGGCA GCA CGAGAGATCTGAATCAGG 365
362 GCGCCCTGAGCAAGCTGAGCGCTGCTCA TCCGGGAA CGGACAGAGCTTGCAGACCTACAG 421
366 GCGCCCTGAGCAAGCTGAGCGCTGCTCA TCCGGGAA CGGACAGAGCTTGCAGACCTACAG 425
422 CGAGCATGGGCAAGCTGACAGAGAGCTCACCAAGACCAC AGCCAGAGCAATTAGAA 481
426 CGAGCATGGGCAAGCTGACAGAGAGCTCACCAAGACCAC AGCCAGAGCAATTAGAA 485
482 GCTGAAGAGCGGTA CGAGCTTGCGCA CGGGGCAAGTGC CCAAGCCGAGAGTAAACA 541
486 GCTGAAGAGCGGTA CGAGCTTGCGCA CGGGGCAAGTGC CCAAGCCGAGAGTAAACA 545
542 GGAAGCCAGCAAGAGCAAGAGACGCTGACAAAGCCAA GAGATATGTGGGAGCCCTGTG 601
546 GGAAGCCAGCAAGAGCAAGAGACGCTGACAAAGCCAA GAGATATGTGGGAGCCCTGTG 605
602 GAACTCTTTGCTCA CCAACCGCTATGTGTGGGCGGTG GGGGCTGCGCAAGCTACCA 661
606 GAACTCTTTGCTCA CCAACCGCTATGTGTGGGCGGTG GGGGCTGCGCAAGCTACCA 665
662 CCAGCA CCA CCAAGCTCTGTGCGCGGCTGCGGCTGCA CTGCAAGAGCACTTGACGA 721
666 CCAGCA CCA CCAAGCTCTGTGCGCGGCTGCGGCTGCA CTGCAAGAGCACTTGACGA 725
722 GGAAGATGGCTTGATCTGAAAGAGATCTGACAGAAATC CTGAGATTAAGACGCTGGT 781
726 GGAAGATGGCTTGATCTGAAAGAGATCTGACAGAAATC CTGAGATTAAGACGCTGGT 785
782 GCAAGATGAGTGTGGCCATTCA CCGGAGAGATGCTGCA GCTGCTGCGGACCTCAAGCC 841
786 GCAAGATGAGTGTGGCCATTCA CCGGAGAGATGCTGCA GCTGCTGCGGACCTCAAGCC 845
842 TGAGGCTGAGTACCAAGGCTTCTGAGAGAGTATGGGTCC GCA CTTGACGTCCACCCCTG 901
846 TGAGGCTGAGTACCAAGGCTTCTGAGAGAGTATGGGTCC GCA CTTGACGTCCACCCCTG 905
902 TGTCAAGTTCGATGATGATCA CTGCTTGAGAGAGGTGA AC CGCTGAGACCTTGAGGAGCTCA 961
906 TGTCAAGTTCGATGATGATCA CTGCTTGAGAGAGGTGA AC CGCTGAGACCTTGAGGAGCTCA 965
962 GCTTGAACGAGCTGATCTGTGAGAGCGTGCAGCA CACGCTGACCTCA GTGACAGATGAGCT 1021
966 GCTTGAACGAGCTGATCTGTGAGAGCGTGCAGCA CACGCTGACCTCA GTGACAGATGAGCT 1025
1022 GGTGTGGGCA CCGAATGATGTGTTCA GGGGGGCAAGAGATGTGTTA CCGAGCTGCAACAGGA 1081
1026 GGTGTGGGCA CCGAATGATGTGTTCA GGGGGGCAAGAGATGTGTTA CCGAGCTGCAACAGGA 1085
1082 GCTCCGGAATGAAGAGAGAACACCCACCCCGGAGCGGATG CACTGCTGGGCAAGAG 1141
1086 GCTCCGGAATGAAGAGAGAACACCCACCCCGGAGCGGATG CACTGCTGGGCAAGAG 1145
1142 GCAAGTGTCTCAAGAGAGACTGCAAGGAGGCTGAGGATG CGCTGTGACCGCCAGGCT 1201
1146 GCAAGTGTCTCAAGAGAGACTGCAAGGAGGCTGAGGATG CGCTGTGACCGCCAGGCT 1205
1202 GCAAGGCGGAGAGAGATTTGCTGAGACCAAGCTGAGAC CACTGAGGCCCGGAGGAGCCCC 1261

1206 GCAAGGCGGAGAGAGATTTGCTGAGACCAAGCTGAGAC CACTGAGGCCCGGAGGAGCCCC 1265
1262 GCTGTGTGCTCTCTGACAGAGATGACCGCACTTCA CAGTGTCTCTGAGACAGAGAGGAGA 1321
1266 GCTGTGTGCTCTCTGACAGAGATGACCGCACTTCA CAGTGTCTCTGAGAGAGAGAGGAGA 1325
1322 GGGGGGAAGAGACACCAAGCTGAGAGATCTTTAAGAGCC CATTCTGAGAAATTTCCGCC 1381
1326 GGGGGGAAGAGACACCAAGCTGAGAGATCTTTAAGAGCC CATTCTGAGAAATTTCCGCC 1385
1382 CAAATTCG----- 1391
1386 CAAATTCGCTCTCTCTCA CCGCTGACGCTCATCTCGAGATG CAGAAAGCCCTGCATGA 1445
1392 ----- 1391
1446 GCAAGTGTGATCA CAGGGGCAATCCGAGGGGAGAGTGTG CTGAGCTGAGTCACTC 1505
1392 ----- 1391
1506 TGGGCACTTCTGTGCGGAGAGACCAAGGCAAGAGATAC TGTGCGGTGCTGTG 1565
1392 ----- 1391
1566 GGAATGTCTGCCCCGGCACTTCA TCAATCCAATCTTGATTA CTTGTACCTGTAGAGAG 1625
1412 GGAAGCTTTCTAGCAATTCCTTGTCTCATGCA CCACTTACAGAGCAAGCCAGAGCCCT 1471
1626 GGAAGCTTTCTAGCAATTCCTTGTCTCATGCA CCACTTACAGAGCAAGCCAGAGCCCT 1685
1472 CACCAAGAGAGTGTGTGTCTGTCACAGGCTGTG CCAAGAGACAGTGGTCTGAA 1531
1686 CACCAAGAGAGTGTGTGTCTGTCACAGGCTGTG CCAAGAGACAGTGGTCTGAA 1745
1532 CCATGAGGACCTGTGTGTGTGTGAGAGATTTGGA CCGGGGAACTTTGGGAGATTTCA 1591
1746 CCATGAGGACCTGTGTGTGTGTGAGAGATTTGGA CCGGGGAACTTTGGGAGATTTCA 1805
1592 CGGACGCTGCGAGCGGACCA CCAACCTGTGTGGCGGTGA AGTCTTGTGAGAGACGCTGCC 1651
1806 CGGACGCTGCGAGCGGACCA CCAACCTGTGTGGCGGTGA AGTCTTGTGAGAGACGCTGCC 1865
1652 ACCTGACCTCAAGGCGCAATTTCTACAGAGAGCGAGAGAT CTTGAGACAGTACAGCCACC 1711
1866 ACCTGACCTCAAGGCGCAATTTCTACAGAGAGCGAGAGAT CTTGAGACAGTACAGCCACC 1925
1712 CAAACATGAGGCGTCAATTGTGTGTGTCAC CCAAGAGAGAGCCCATCTACATGCTGATGA 1771
1926 CAAACATGAGGCGTCAATTGTGTGTGTCAC CCAAGAGAGAGCCCATCTACATGCTGATGA 1985
1772 GCTTGTGAGGGGGGCACTTCTGACCTTCTCCGCA CCGAGGGGGCCGCTGCGGGT 1831
1986 GCTTGTGAGGGGGGCACTTCTGACCTTCTCCGCA CCGAGGGGGCCGCTGCGGGT 2045
1832 GAAAGCTTGTGTCAGATGTGTGGGAGATGCAAGCTG CCGCATGAGATCTGAGAGCA 1891
2046 GAAAGCTTGTGTCAGATGTGTGGGAGATGCAAGCTG CCGCATGAGATCTGAGAGCA 2105
1892 GTGTGTCACCA CCGGGAAGCTGTGTCGGAAC TGTGCTGTGAGACAGAGAAATGCTCT 1951
2106 GTGTGTCACCA CCGGGAAGCTGTGTCGGAAC TGTGCTGTGAGACAGAGAAATGCTCT 2165
1952 GAAAGTCAAGTGACTTTGGAGTATCCGAGAGAGAGCGA TGGGGTCTATGACAGCTCAGG 2011
2166 GAAAGTCAAGTGACTTTGGAGTATCCGAGAGAGAGCGA TGGGGTCTATGACAGCTCAGG 2225
2012 GGGCTTCAAGCAAGTCTCCCTGAGATGGAAC CCGCACTGAGGCGCTTAACTACGCGCTGA 2071
2226 GGGCTTCAAGCAAGTCTCCCTGAGATGGAAC CCGCACTGAGGCGCTTAACTACGCGCTGA 2285
2072 CTCTCCGAAGCGAGCTGTGAGCTTTGGACATCTTTGCT GTGGGAGACCTTCAAGCTTGG 2131

Db	2286	CTCCTCCGAAGGCGACGCTGTGGAGCTTTGGCATCTTGTCTTGGAGACCTTCAGCTGGG	23458
Qy	2132	GGCCTCCCTATCCCAACTTCAGCAATCAGACACGCGAGATTGTGAGAAAGGGGGG	21911
Db	2346	GGCCTCCCCCTATCCCAACTCAGCAATCAGACACACGCGAGATTGTGAGAAAGGGGGG	24050
Qy	2192	CCGCTGCGCCCTGCCAGAGCTGTGTCTCGATGCGCGTGTTCAGGCTCAGAGACAGTGGTG	22513
Db	2406	CCGCTGCGCCCTGCCAGAGCTGTGTCTCGATGCGCGTGTTCAGGCTCAGAGACAGTGGTG	24655
Qy	2252	GGCCTATGAGCCTGGGCGACGGGCCAGGCTTCAGACACATCTTACAGAGAGCTGCAGAGAT	23111
Db	2466	GGCCTATGAGCCTGGGCGACGGGCCAGGCTTCAGACACATCTTACAGAGAGCTGCAGAGAT	25253
Qy	2312	CGAAGGCGGATCGGTGAGGCTGGGACCCCTTCTCAAGCTGTGTGCTTCGACGGCT	23711
Db	2526	CGAAGGCGGATCGGTGAGGCTGGGACCCCTTCTCAAGCTGTGTGCTTCGACGGCT	25858
Qy	2372	AGGTGCACTCTTCAGCGGGCTTCAGACTCATATGCTGACAGCTCTTCAACAGTCTGGAATC	24311
Db	2586	AGGTGCACTCTTCAGCGGGCTTCAGACTCATATGCTGACAGCTCTTCAACAGTCTGGAATC	26458
Qy	2432	CTGCCACGAGATCCACACTGCGCGGACAGATGACGCGCGTGTCTCTGTGCTCCCTGC	24919
Db	2646	CTGCCACGAGATCCACACTGCGCGGACAGATGACGCGCGTGTCTCTGTGCTCCCTGC	27050
Qy	2492	TGCTGCCAGGGCTTCTCTTCCTCGGCGAGAAACAAATMAAACCATTTGTCGCACTGAAAA	25511
Db	2706	TGCTGCCAGGGCTTCTCTTCCTCGGCGAGAAACAAATMAAACCATTTGTCGCACTGAAAA	27655
Qy	2552	AAAAAA 2557	
Db	2766	AAAAAA 2771	
RESULT 5			
AX281847			
LOCUS	AX281847	2889 bp	DNA linear PAT 02-NOV-2001
DEFINITION	Sequence 256 from Patent WO0177389.		
ACCESSION	AX281847		
VERSION	AX281847.1	GI:16609098	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE			
AUTHORS	Shiffman,D., Somogyi,R., Lawn,R., Sellhammer,J.J., Porter,G.J., Mikhla,T. and Tai,J.		
TITLE	Genes expressed in foam cell differentiation		
JOURNAL	Patent: WO 0177389-A 256 18-OCT-2001;		
FEATURES	Incyte Genomics, Inc. (US)		
source	Location/Qualifiers		
	1..2889		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: 997347.6"		
ORIGIN			
Query Match	87.0%;	Score 2326.4;	DB 6; Length 2889;
Best Local Similarity	92.3%;	Pred. No. 0;	
Matches 2558; Conservative	0;	Mismatches 1;	Indels 211; Gaps 2;
Qy	2	CCGGGCTCCGACCCGGGCTGAGTGGTCCGAGGGCCGTCGCCAGAGACAGTCCCGCTGCG	61
Db	118	CCGGGCTCCGACCCGGGCTGAGTGGTCCGAGGGCCGTCGCCAGAGACAGTCCCGCTGCG	177
Qy	62	GAACGACATATGGGCTTCTTTCTGAGCTGTGCAGCCGCCAGGGGACAGGGGATCTGTGA	121
Db	178	GAACGACATATGGGCTTCTTTCTGAGCTGTGCAGCCGCCAGGGGACAGGGGATCTGTGA	237
Qy	122	GCAATGACAGAGCCGAGCTTGTCTTACTGAGAGGGCATGAGAAAGTGAATGGCCACG	181

Db	238	GCAATGCAAGAGGCCCAAGCTTGGTCTAATCTGAGGGCATGAGAAATGTGATGGCCACGG	297
Qy	182	GGTCAGAAGTGCAGGGAGTATGCAAGACTGTCTTACCACATGTCCCTGCAGACAGTGG	241
Db	298	GGTCAAGATGACAGGAGTATGCAAGACTGTCTTACCACATGTCCCTGCAGACAGTGG	357
Qy	242	GGGCGAAGCCGGGSCCTACAGCCTTGACAGCCCATATGATCATGTCTTGGGTGATGATC	301
Db	358	GGGCGAAGCCGGGSCCTACAGCCTTGACAGCCCATATGATCATGTCTTGGGTGATGATC	417
Qy	302	CAGCCAAACTGA--GGGCTGAGCGCTTGCTGCGGACAGCGCAGAGSATTCTGAATCAG	360
Db	418	CAGCCAAACTGAGGGGCTGAGCGGCTTGTCTGCGGACAGCAGAGATCTGAATCTCAG	477
Qy	361	GGCCCTTGAGCAGACTGAGCCTGTCTCATTCGGGAGCGGACAGCCTTTCGACAGCTTACA	420
Db	478	GGCCCTTGAGCAGACTGAGCCTGTCTCATTCGGGAGCGGACAGCCTTTCGACAGCTTACA	537
Qy	421	CGAGCAGTGGCAGACAGCTGCAGCAGAGACTTACCAAGATCCCAAGCCAGACATTGAGA	480
Db	538	CGAGCAGTGGCAGACAGCTGCAGCAGAGACTTACCAAGATCCCAAGCCAGACATTGAGA	597
Qy	481	AGCTGAAGAGCGCAATACCGAGCTGTGGACAAGGACAGTGGCCAAAGCCAGCGCAAGTACC	540
Db	598	AGCTGAAGAGCGCAATACCGAGCTTGTGGACAAGGACAGTGGCCAAAGCCAGCGCAAGTACC	657
Qy	541	AGGAGGCGCAGCAAGACAGAGACCGGTGACAAAGCCAAAGACAAAGTATGTGGACCGCTGT	600
Db	658	AGGAGGCGCAGCAAGACAGAGACCGGTGACAAAGCCAAAGACAAAGTATGTACGCGCCTGT	717
Qy	601	GGAAAGCTTTTGTCTACCAACAACCGCTATGTGTCTGGCGTTCGGGCTTCGCAAGTAAAC	660
Db	718	GGAAAGCTTTTGTCTACCAACAACCGCTATGTGTCTGGCGTTCGGGCTTCGCAAGTAAAC	777
Qy	661	ACCAAGCACCAACACAGCTCCTGTGCGCCGGGCTGTGCGGATCATCTGACAGGACCTGCACG	720
Db	778	ACCAAGCACCAACACAGCTCCTGTGCGCCGGGCTGTGCGGATCATCTGACAGGACCTGCACG	837
Qy	721	AGGAGATGGCTTGCATCCTGAAAGAGATTCCTGCAGGAATACCTGAGATTAAGACCTGG	780
Db	838	AGGAGATGGCTTGCATCCTGAAAGAGATTCCTGCAGGAATACCTGAGATTAAGACCTGG	897
Qy	781	TGCAGATAGAGTGTGTGGCCATTACCGGGAGATGGCTGCAGCTGTCTGCCGATTCACAG	840
Db	898	TGCAGATAGAGTGTGTGGCCATTACCGGGAGATGGCTGCAGCTGTCTGCCGATTCACAG	957
Qy	841	CTGAGGCGTGAATCAAGGCGTTCCTGCGACAGATATGGGTCCGCACTGACGTCCACCCCT	900
Db	958	CTGAGGCGTGAATCAAGGCGTTCCTGCGACAGATATGGGTCCGCACTGACGTCCACCCCT	1017
Qy	901	GTCATCAGCTTCGATGATGATCATCTGCTTGAGAGAGGTGAAACCGCTGAGCCTTGGGAGCTCC	960
Db	1018	GTCATCAGCTTCGATGATGATCATCTGCTTGAGAGAGGTGAAACCGCTGAGCCTTGGGAGCTCC	1077
Qy	961	AGCTGAACGAGCTTACTGTGAGAGCGGTGCACACACGCTGACTCATGTGACAGATGAGC	1020
Db	1078	AGCTGAACGAGCTTACTGTGAGAGCGGTGCACACACGCTGACTCATGTGACAGATGAGC	1137
Qy	1021	TGGCTGTGGCCACCGAGATGGTGTTCAGGCGGCGAGAGATGGTTATCGAGCTGCAACAGG	1080
Db	1138	TGGCTGTGGCCACCGAGATGGTGTTCAGGCGGCGAGAGATGGTTATCGAGCTGCAACAGG	1197
Qy	1081	AGCTCCGGAATGAAGAGAGAACCAACCCCGGAGCGGGTCACTGCTTGGGCAAGA	1140
Db	1198	AGCTCCGGAATGAAGAGAGAACCAACCCCGGAGCGGGTCACTGCTTGGGCAAGA	1255
Qy	1141	GGCAAGTGTCTCAAGAAAGCACTGCAGGGGCGTGCAGGTAGCGCTGTGCAGCCAGGCAAGC	1200
Db	1258	GGCAAGTGTCTCAAGAAAGCACTGCAGGGGCGTGCAGGTAGCGCTGTGCAGCCAGGCAAGC	1317
Qy	1201	TGCAGGCCACAGAGAGTGTGCTGCAGACCAAGCTGAGACACTGGGCCCCGGGCCAGCCCC	1266

Db 1318 TGCAGGCCAGAGAGATTCTGACAGCAAGCTGGAGCACTGGGCCCCGGGAGAGCCCC 1377
 Qy 1261 CGCTTGTGCTGCTCTCTGACAGATGATCCGCACTCCAGTGTCTCTGAGACAGAGAGAG 1320
 Db 1378 CGCTGTGCTGCTCTCTGACAGATGATCCGCACTCCAGTGTCTCTGAGAGAGAGAGAG 1437
 Qy 1321 AGGGGGGAGAGAGACCCAGCTGGAGATCTTAAAGGAGCAATCTAGGAATCTTCGGCC 1380
 Db 1438 AGGGGGGAGAGAGACCCAGCTGGAGATCTTAAAGGAGCAATCTAGGAATCTTCGGCC 1497
 Qy 1381 CCAAGTTCTG----- 1391
 Db 1498 CCAAGTTCTGCGTCCCTCCAGCGCTGACAGTCAATCCGAGAGTGGAGAAAGCCCTGCAATG 1557
 Qy 1392 ----- 1391
 Db 1558 AGCAGCTGTGTATCCAGGGGGCCATCCGAGGGCAGAGGTGGCTGAGCTGTGTGCACT 1617
 Qy 1392 ----- 1391
 Db 1618 CTGGGAGCTTCTCGTGTGGGAGAGCCAGGGCAGAGAGATCGTGTGTGTGCTGT 1677
 Qy 1392 -----AACGTGATCCGACTGGAG 1410
 Db 1678 GGGATGTGTGCCCCGGGCACTTCATGATCAAGTCTTGGATTAACCTGTACCGACTGGAG 1737
 Qy 1411 GGGAGGCTTCTCTGACATTCCTTGTCTATGACCACTACTGAGACCCAGAGCCCC 1470
 Db 1738 GGGAGGCTTCTCTGACATTCCTTGTCTATGACCACTACTGAGACCCAGAGCCCC 1797
 Qy 1471 TCACCAAGAGAGAGTGTGTGTCTGACAGAGGCTGTGCCAAGAGCAATGTGGTCTGA 1530
 Db 1798 TCACCAAGAGAGAGTGTGTGTCTGACAGAGGCTGTGCCAAGAGCAATGTGGTCTGA 1857
 Qy 1531 ACCATGAGAGACCTGT 1590
 Db 1858 ACCATGAGAGACCTGT 1917
 Qy 1591 GGGAGGCTTCTGAGAGCCGACCAACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1650
 Db 1918 GGGAGGCTTCTGAGAGCCGACCAACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1977
 Qy 1651 GACCTGACCTCAAGGCGCAAGTTCTTACAGAGAGCGAGATCTTGAAGCGATCAAGCAAC 1710
 Db 1978 GACCTGACCTCAAGGCGCAAGTTCTTACAGAGAGCGAGATCTTGAAGCGATCAAGCAAC 2037
 Qy 1711 CCAACATCGT 1770
 Db 2038 CCAACATCGT 2097
 Qy 1771 AGCTTGTGAGGGGGGCGACTTCTGTGACCTTCTCTGCGACGAGAGGGGGCCGCTGTGGG 1830
 Db 2098 AGCTTGTGAGGGGGGCGACTTCTGTGACCTTCTCTGCGACGAGAGGGGGCCGCTGTGGG 2157
 Qy 1831 TGAAGATCTGT 1890
 Db 2158 TGAAGATCTGT 2217
 Qy 1891 AGTGTGATCCAGCGGAGCTGT 1950
 Db 2218 AGTGTGATCCAGCGGAGCTGT 2277
 Qy 1951 TGAAGATCAATGATCTTGGAGATGTCCGAGAGAGAGCCGATGGGGGTCTATGAGCCTCAG 2010
 Db 2278 TGAAGATCAATGATCTTGGAGATGTCCGAGAGAGAGCCGATGGGGGTCTATGAGCCTCAG 2337
 Qy 2011 GGGGGCTGAGAGAGTCCCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2070
 Db 2338 GGGGGCTGAGAGAGTCCCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2397
 Qy 2071 ACTCTCCGAGAGAGAGCTGT 2130
 Db 2398 ACTCTCCGAGAGAGAGCTGT 2457

Qy 2131 GGGCTCTCCCTCTATCCCACTTCAGCATACAGACAGAGAGTTGTGGAGAGAGGGG 2190
 Db 2458 GGGCTCTCCCTCTATCCCACTTCAGCATACAGACAGAGAGTTGTGGAGAGAGGGG 2517
 Qy 2191 GCGGTGTGCTGCGGAGAGCTGT 2250
 Db 2518 GCGGTGTGCTGCGGAGAGCTGT 2577
 Qy 2251 GGGCTATAGAGCTGTGGAGCGGCGCCAGCTTACAGACCATATACAGAGAGCTGACAGCA 2310
 Db 2578 GGGCTATAGAGCTGTGGAGCGGCGCCAGCTTACAGACCATATACAGAGAGCTGACAGCA 2637
 Qy 2311 TCCGAAAGCGGATGT 2370
 Db 2638 TCCGAAAGCGGATGT 2697
 Qy 2371 TAGGTGACAGCTCTGAGCGGCTCCAGCTCATATAGCTGACAGCTTTCACAGTCTGAGCT 2430
 Db 2698 TAGGTGACAGCTCTGAGCGGCTCCAGCTCATATAGCTGACAGCTTTCACAGTCTGAGCT 2757
 Qy 2431 CCTGCCACAGATCCACACTGCGCGAGATGACAGCGCGGTCTCTGTGTGCTGT 2490
 Db 2758 CCTGCCACAGATCCACACTGCGCGAGATGACAGCGCGGTCTCTGTGTGCTGT 2817
 Qy 2491 CTGTGCTGCGGCTCTCTCTTCCGGGAGAGAAACATTAACCACTTGTGCTGAGAA 2550
 Db 2818 CTGTGCTGCGGCTCTCTCTTCCGGGAGAGAAACATTAACCACTTGTGCTGAGAA 2877
 Qy 2551 AAAAAAAAAA 2560
 Db 2878 AAAAAAAAAA 2887

RESULT 6
 LOCUS CQ725220 2697 bp DNA linear PART 03-FEB-2004
 DEFINITION Sequence 11154 from Patent WO2068579.
 VERSION CQ725220
 KEYWORDS CQ725220.1 GI:42286077
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 TITLE Kites, such as nucleic acid arrays, comprising a majority of
 humexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 11154 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES
 source 1..2697
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN

Query Match 84.8%; Score 2267; DB 6; Length 2697;
 Best Local Similarity 92.2%; Pred. No. 0; Mismatches 0; Indels 210; Gaps 1;
 Matches 2487; Conservative 0; Mismatches 0; Indels 210; Gaps 1;
 Qy 62 GAACAGCACTATGGCTTCTTCTGTAGAGTGTGAGCCCGGAGGAGCAAGGGGTCTGTGA 121
 Db 1 GAACAGCACTATGGCTTCTTCTGTAGAGTGTGAGCCCGGAGGAGCAAGGGGTCTGTGA 60
 Qy 122 GCAATGACAGAGCGGAGCTTCTGTACTGTGAGGAGCATGAGAAAGTGATGGCCAGCG 181
 Db 61 GCAATGACAGAGCGGAGCTTCTGTACTGTGAGGAGCATGAGAAAGTGATGGCCAGCG 120
 Qy 182 GGTCAAGGTGACAGGAGTATGAGGACTGCTTACCAACATGTCCCTGACAGAGTGG 241

Db	121	GGTCAAGATGTACAGGAGATGTGACAGACTGTCTTACCAACATGTTCCTCGACGACGACATGG	180
Qy	242	GGGCGAAGGCCGGGSCCATCAGCCCTTGAACGCCCATCATGTCAGTCTTGGGCTGATGATCAC	301
Db	181	GGGCGACAGCGGGGSCCATCAGCCCTTGAACGCCCATCATGTCAGTCTTGGGCTGATGATCAC	240
Qy	302	CAGCCAAACTGAGGGGSCCTGAGCCGCTTGGTCGGGACAGACGACAAAGATCTGAACTCAG	361
Db	241	CAGCCAAACTGAGGGGSCCTGAGCCGCTTGGTCGGGACAGACGACAAAGATCTGAACTCAG	300
Qy	362	GGCCCTGAGCAAGGCTGAGCCTGTCTCATCCGGGAAACGGCAGAGGTTGGCAAGACTACAG	421
Db	301	GGCCCTGAGCAAGGCTGAGCCTGTCTCATCCGGGAAACGGCAGAGGTTGGCAAGACTACAG	360
Qy	422	CGACCAAGTGGCAGCAGCTTGACGACGAGAGCTCACCAAGCCCAACGCCAAGTACCA	481
Db	361	CGACCAAGTGGCAGCAGCTTGACGACGAGAGCTCACCAAGCCCAACGCCAAGTACCA	420
Qy	482	GCTAAAGGCCAAGTACCGAGGCTCTGGCACCGGGACAGTGGCCCAAGCCAAAGCCAAAGTACCA	541
Db	421	GCTAAAGGCCAAGTACCGAGGCTCTGGCACCGGGACAGTGGCCCAAGCCAAAGCCAAAGTACCA	480
Qy	542	GGAGGCCAGCAAAAGACAAGACCCGTGACAAAGGCCAAAGACAAGATATGAGCCGACCTGTG	601
Db	481	GGAGGCCAGCAAAAGACAAGACCCGTGACAAAGGCCAAAGACAAGATATGAGCCGACCTGTG	540
Qy	602	GAACTCTTTTGTCTCAACAACCGCTATGTGTGGGCGTGGGCGCTGGCAGCTACACA	661
Db	541	GAACTCTTTTGTCTCAACAACCGCTATGTGTGGGCGTGGGCGCTGGCAGCTACACA	600
Qy	662	CCAGCACACACACACAGCTCTTGTGTGCCCCGGCTGTGTGGTCACTGACAGACCTTGACAGA	721
Db	601	CCAGCACACACACACAGCTCTTGTGTGCCCCGGCTGTGTGGTCACTGACAGACCTTGACAGA	660
Qy	722	GGAGTGGCTTGACCTCTGAAAGAGATCTGTACAGAAATACCTGGAGATTTAGCAGCTGGT	781
Db	661	GGAGTGGCTTGACCTCTGAAAGAGATCTGTGAAAGAGATCTGTGAAAGATTTAGCAGCTGGT	720
Qy	782	GCAGAGATGAGGTGTGTGGCATTCAACCGGAGATGTGCTGACGCTGTCCGCGATCCAGCC	841
Db	721	GCAGAGATGAGGTGTGTGGCATTCAACCGGAGATGTGCTGACGCTGTCCGCGATCCAGCC	780
Qy	842	TGAGGCTGAGTACCAAGGCTTCTTGCGCACATATATGGTTCGGCACCTGACGTGCCACCTGT	901
Db	781	TGAGGCTGAGTACCAAGGCTTCTTGCGCACATATATGGTTCGGCACCTGACGTGCCACCTGT	840
Qy	902	TGTCACTGTCATGTGATGTCATGCTTTGAGAGAGGTGAACCGGTGGAGCCTGGGAGACTTCA	961
Db	841	TGTCACTGTCATGTGATGTCATGCTTTGAGAGAGGTGAACCGGTGGAGCCTGGGAGACTTCA	900
Qy	962	GCTGACAGAGCTGACTGTGTGAGAGGCTGACACACACGCTGACCTCAGTGAACATATGACT	1021
Db	901	GCTGACAGAGCTGACTGTGTGAGAGGCTGACACACACGCTGACCTCAGTGAACATATGACT	960
Qy	1022	GACTGTGGCCACCGAGATGTGTTCAGGGCGGACAGAGATGTGTTACGACGCTGCAACAGA	1081
Db	961	GACTGTGGCCACCGAGATGTGTTCAGGGCGGACAGAGATGTGTTACGACGCTGCAACAGA	1020
Qy	1082	GCTCCGGAATGAAGAGAGAACCCACCCCGGGAGGGGTGGACGCTGTGGGCAAGAG	1141
Db	1021	GCTCCGGAATGAAGAGAGAACCCACCCCGGGAGGGGTGGACGCTGTGGGCAAGAG	1080
Qy	1142	GCAAGTGTGCAAGAGACCTGACAGGGGCTGACAGTATCGCTGTGCAAGCCAGGCCAAAGCT	1201
Db	1081	GCAAGTGTGCAAGAGACCTGACAGGGGCTGACAGTATCGCTGTGCAAGCCAGGCCAAAGCT	1140
Qy	1202	GCAGGCCACAGAGAGATTGTGACACCAAGCTGAGACCTTGGGCCCCGGCGAGCCCCC	1261
Db	1141	GCAGGCCACAGAGAGATTGTGACACCAAGCTGAGACCTTGGGCCCCGGCGAGCCCCC	1200
Qy	1262	GCTGTGTGCTCTCTCTGACAGATGACCCGCACTTCACTGTGTCTCTGAGACGAGACGACAGA	1321
Db	1201	GCTGTGTGCTCTCTCTGACAGATGACCCGCACTTCACTGTGTCTCTGAGACGAGACGACAGA	1260

QY	1332	GGGGGGGAGGACACCCAGCGTGGAGATCCTTTAAAGGCGCAATCTGAGGAATCTTCGGCC	1391
Db	1261	GGGGGGAAGGACACCCAGCGTGGAGATCCTTTAAAGGCGCAATCTGAGGAATCTTCGGCC	1320
QY	1382	CAAGTTCG-----	1391
Db	1391	CAAGTTCGCGTCCCTCCACCGCTGACGCTATTCGGAGGTGACAGAGCCCTGCATGA	1380
QY	1392	-----	1391
Db	1381	GCAAGCTGTGTACACACGGGGCCATCCGAGGGCAGAGGTGGCTGACGTGCTGCACTC	1440
QY	1392	-----	1391
Db	1441	TGGGGACTTCCTGTGTGGGGAGGACAGGGCAAGACAGATGCTGTGCTGTGCTGTG	1500
QY	1392	-----	1391
Db	1501	GGATGCTGCGCCCGGCACTTCATCATCCAGTCTTTGATTAACCTGTACCGATGTGAAG	1560
QY	1412	GGAAAGCTTTCCTAGCATTTCTTTGCTGATCGACCACTACTAGAGCACCGAGCCCT	1471
Db	1561	GGAAAGCTTTCCTAGCATTTCTTTGCTGATCGACCACTACTAGAGCACCGAGCCCT	1520
QY	1472	CACCAAGAAGTGTGTGTCTTGCACAGGGCTGTGCCAAGGACAAAGTGGTGTGAA	1531
Db	1621	CACCAAGAAGTGTGTGTCTTGCACAGGGCTGTGCCAAGGACAAAGTGGTGTGAA	1680
QY	1532	CCATAGAGACCTGTGTGTGGGTAGACAGATTTAGACGGGGGAACTTTGGGAGATGTTCA	1591
Db	1681	CCATAGAGACCTGTGTGTGGGTAGACAGATTTAGACGGGGGAACTTTGGGAGATGTTCA	1740
QY	1592	CGGAGCGCTGCGAGCGCGACAACACCGCTGTGGCGGTGAAGTCTTGTGAGAGACGCTCC	1651
Db	1741	CGGAGCGCTGCGAGCGCGACAACACCGCTGTGGCGGTGAAGTCTTGTGAGAGACGCTCC	1800
QY	1652	ACCTGACCTCAAGGCGCAAGTTCTACAGAGACGAGAGATCTGAAACAGTACAGCCACC	1711
Db	1801	ACCTGACCTCAAGGCGCAAGTTCTACAGAGACGAGAGATCTGAAACAGTACAGCCACC	1860
QY	1712	CAACATCGNGCGTCAATTTGTGTGCAACCCAGAAAGACCCATCTACATGCTCATGA	1771
Db	1861	CAACATCGNGCGTCAATTTGTGTGCAACCCAGAAAGACCCATCTACATGCTCATGA	1920
QY	1772	GCTTGTGACAGGGGGCGAATTCCTGAACCTTCTCGCACGAGAGGGGCGCGCTCGGGGT	1831
Db	1921	GCTTGTGACAGGGGGCGAATTCCTGAACCTTCTCGCACGAGAGGGGCGCGCTCGGGGT	1980
QY	1832	GAAAGCTCTGCTGAGATGTGTGGGGATGTGCACTGTGTGCAATGAGATCACTGGAGAGCA	1891
Db	1981	GAAAGCTCTGCTGAGATGTGTGGGGATGTGCACTGTGTGCAATGAGATCACTGGAGAGCA	2040
QY	1892	GTGCTGCATCCACCGGGACCTTGGCTTCGGAATCTGCTGTGTGACAGAGAGATGTCT	1951
Db	2041	GTGCTGCATCCACCGGGACCTTGGCTTCGGAATCTGCTGTGTGACAGAGAGATGTCT	2100
QY	1952	GAAATCAATGATCTTTGGGATGTCCCGAGAGAGAGCCGATGGGGTCTATGCAACCTCAGG	2011
Db	2101	GAAATCAATGATCTTTGGGATGTCCCGAGAGAGAGCCGATGGGGTCTATGCAACCTCAGG	2160
QY	2012	GGGCTCAGACAAAGTCCCGGTGAGGTGACCGCACTGAGGCCCTTAACCTACGGCGCTA	2071
Db	2161	GGGCTCAGACAAAGTCCCGGTGAGGTGACCGCACTGAGGCCCTTAACCTACGGCGCTA	2220
QY	2072	CTCCTCGAAGACGACGTGTGAGGCTTTGGCATCTTGTCTGTGGAGACCTTCAGCCTGGG	2131
Db	2221	CTCCTCGAAGACGACGTGTGAGGCTTTGGCATCTTGTCTGTGGAGACCTTCAGCCTGGG	2280
QY	2132	GGCTCCTCCCTATCCCAACCTTAGCAATCAGACAGACGGGAGTTGTGGAGAAAGGGGGG	2191
Db	2281	GGCTCCTCCCTATCCCAACCTTAGCAATCAGACAGACGGGAGTTGTGGAGAAAGGGGGG	2340

Qy	2192	CCGCTCGCCCTGCCCAAGCTGTGTCTCTGATGCCGTGTTCAAGGCTCAATGAGACAGTGGT	2253
Db	2341	CCGCTCGCCCTGCCCAAGCTGTGTCTCTGATGCCGTGTTCAAGGCTCAATGAGACAGTGGT	2400
Qy	2252	GGCCTATGAGCTGGGCGAGCGGCCAGCTTACAGCACATCTACACAGAGCTGCGAGCAT	2311
Db	2401	GGCCTATGAGCTGGGCGAGCGGCCAGCTTACAGCACATCTACACAGAGCTGCGAGCAT	2460
Qy	2312	CCGAAAGGGGCGATCGGTGAGGCTGTGGACCCCTTCTCAAGCTGTGTGCTTCGACAGGCT	2371
Db	2461	CCGAAAGGGGCGATCGGTGAGGCTGTGGACCCCTTCTCAAGCTGTGTGCTTCGACAGGCT	2520
Qy	2372	AGGTGCAGCTCTCAGCGGCTCCAGCTCATATGCTGACAGCTCTTCAACAGTCTTGAATTC	2431
Db	2521	AGGTGCAGCTCTCAGCGGCTCCAGCTCATATGCTGACAGCTCTTCAACAGTCTTGAATTC	2580
Qy	2432	CTGCCACAGCATCACTGCGCGGACAGATGACAGCGCGGTCTCTCTGTGTCCCTGC	2491
Db	2581	CTGCCACAGCATCACTGCGCGGACAGATGACAGCGCGGTCTCTCTGTGTCCCTGC	2640
Qy	2492	TGCTGCACAGGCTTCTCTTCCGGGCAAGAACATATAAACACACTGTGCCACTGAA	2548
Db	2641	TGCTGCACAGGCTTCTCTTCCGGGCAAGAACATATAAACACACTGTGCCACTGAA	2697

RESULT 7

MPRESSCR				
LOCUS	MMFESCR	2680 bp	mRNA	linear
DEFINITION	Mouse c-fes proto-oncogene mRNA for c-fes protein.			ROD 05-JUN-1995
ACCESSION	X12616			
VERSION	X12616.1	GI:50955		
KEYWORDS	fes cellular oncogene; fes proto-oncogene; proto-oncogene; tyrosine kinase.			
SOURCE	Mouse musculus			
ORGANISM	Mus musculus			
REFERENCE	Mammalia; Euteleostomi; Chordata; Vertebrata; Euteleostomi; Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Wilks, A.F. and Kurban, R.R.			
TITLE	Isolation and structural analysis of murine c-fes cDNA clones			
JOURNAL	Oncogene 3 (3), 289-294 (1988)			
MEDLINE	89083198			
PUBMED	3060793			
FEATURES	Location/Qualifiers			
SOURCE	1..2680			
	/organism="Mus musculus"			
	/mol_type="mRNA"			
	/db_xref="taxon:10090"			
	/clone_1lb="lambda gct11."			
	60..2522			
	/note="unnamed protein product; c-fes protein"			
	/codon_start=1			
	/protein_id="CAA3138.1"			
	/db_xref="GI:50956"			
	/db_xref="GDB:P16879"			
	/db_xref="UniProt/Swiss-Prot:P16879"			
	/translation="MGFSESLSPQGHGAVVOQMEABRLLEGKMKMAQKVKSDREYAGLHHMSLDGSGGSMGSDSPVSGSMATITQETLSYLKQHAEDLNSGSLSKSVLIRHNSLRKTYNEMQQLQOELTKHSQDILKTKOYTLVRDSQARKYQKASKDKDRDRAKDKYVRSIMLTPAHNRKRYLGVAQAQLHHNHRPMLPGLSLQDLHEBMGLIKDLDLYELISLVQDDVASIHRELLAARLQPEBYIGFLKQYSTDPVPPCVTDESLDEGEQLPEQLDNLLETLSVOHTLTSTVDLAVATKELASQENVSQLQRELESEBONTHPRERVOLTSKRQMLDEALIGQIALCSQDLQAOELDSKMEQGLTGEPAYVALDDHRSTSTREGGRPTLEILSKSHFGIIRPKPSIIPPLVPEVQKLYEOLMNGAIPMAEVAELLTHTGDPLNBSGQKQYVLSWMDGPHRHTIISLQNLTYLEEDGPSIPLITHLSSOQLPTKSGVTLFRAVPDKNYKHEDVLASQICRGNGEFSGKRLADNTPVAKSCRETIIPDLKAKFLDEALIKQNHNPVTLIGVQTKQKPIYIMVLVQGGFLTLRTGALRVKTLIQWGDAAAGMEYLESKCTHRDIAARCLYETKVKLISDGMSREEDQIVAAKCSGLRQVPVMTAEPALNYGRYSSESVMSGILMETFSGASPYPLTINQOTREFVEKGRHLPCEPLCDPAVRLMEQCVAAYPGQRPSPSIIQDELHSIRKRR"			

Query Match	58.0%;	Score 1552.2;	DB 10;	Length 2680;
Beet Local Similarity	79.1%;	Pred. No. 7,96-228;		
Matches 1970;	Conservative	0;	Mismatches 303;	Indels 216; Gaps 2;
QY	CGGAAACAGCACTATGAGGCTTCTCTTTCAGACTGTGACGCCCGGAGCCCAACGGGGTCTCTG	119		
DB	48 CAGGACACCATATATGAGGCTTCTCTTTCAGAGCTGTGACGCCCGGAGCCCAACGGGGTCTCTG	107		
QY	120 CAGCAAAATGACAGAGGCGGACCTTCCGTCTATCTGAGGGCATGAGAAATGAGATGGCCAG	179		
DB	108 CAGCAGATGACAGGAAAGCTGAGCTGCGCTTATTTGAGAGGGCATGAGAAATGAGATGGCCAG	167		
QY	180 CGGGTCAAGATGACAGGAGATATGACAGACTGCTTTACCAACATATGTCCTTCAGAGAAAGT	239		
DB	168 AGGGTCAAGATGACCGGGAAATATGACAGGATTTGGCTTACCAACATATGTCCTTCAGAGAAAGC	227		
QY	240 GGGGGCCAGAGCCGGGGCCATCAGCCCTGACAGAGCCCACTCATGTCAAGTCTCTGGGCTGAGATC	299		
DB	228 GGAAGCCAGAGCTGAGAGAGAGGAGGAGCCCGGACAGCCCTTCTCAGCCAGTCTTGGGACAGATATA	287		
QY	300 ACCAGCCAAATGAGGGGCTGAGCGCGCTTGTCTGTGGGACAGACGACAGAGATCTTGAATCTCA	359		
DB	288 ACAAGCCAGACCCAGGAATCTGAGCGGGGTGTGGGACAGATGACAGAAATCTGAATCTGAC	347		
QY	360 GGGGCTCTGAGCAAGCTGAGCTGTCTATCTCGGGAAATGGCAGCAGCTTCCGAGAACTTCA	419		
DB	348 GGGGCTCTTGAAGCAAACTGAGGGTGTCTATCCGGGAGCGGACAGCTCTGAGAAATAAGCTGAC	407		
QY	420 AGCAGAGATGGGACGAGCTGACGACGAGAGGTCAACAAAGCCACAGCCAGCAATTTGAG	479		
DB	408 AACAGCAGATGGGACGAGCTGACGACGAGAGGTCTACCAAGACCCACAGCCAGAGCAATTTGAG	467		
QY	480 AAGCTGAAGACCAATGACCGAGCTCTGGCAACGGGACAGTCCCAAGCCAAAGCCAGATAC	539		
DB	468 AAGCTGAAGACTCATGATCCGGAACCTCTGTATGAGATGACACCAAGCCAGCCAGATGAC	527		
QY	540 CAGGAGCCAGCAAAAGCAAGAACCTGTGACAGAGCCAAAGATATGTGCGCAGCTGT	599		
DB	528 CAGGAAACCCAGCAAAAGCAAGATATGAGAAAGGCCAAAGCAAGATATGTGCGGAGCTGT	587		
QY	600 TTGAAAGCTTTTGTCTACCAACACCGGTATGTGTGGGGCGAGGGGCTGGGACGATACAC	659		
DB	588 TTGAAAGCTTTTGTGCCCAACCAACCGGTATGATGTCTGGGTGTGAGGGGCGGACAGCTGAC	647		
QY	660 CACCAAGACACACACACAGCTCTGTGCTGCCGGGCTGTGTGGGTCTCATCTGACAGAACCTGTAC	719		
DB	648 CACCAACACACACACACCGGTTCATGTCTGCTGTGCTGTGACAGTACATGACAGAAATTTGAC	707		
QY	720 GAGGAGATGGCTTGTGCATCTTGAAGAGATCTTGACGAAATATCTGAGATTTAGCAGCTGT	779		
DB	708 GAGGAGATGGCGGGCATTTCTGAAGAGCATCTTGAGAGAAATATCTGAGATTTAGCAGCTGT	767		
QY	780 GTGACAGATGAGGTGTGTGTGCTTATCACCAGGAGATGTGCTGACAGTGTGCCCGGATCTAG	839		
DB	768 GTGACAGACAGATGTGGCATCCATTCACCGTGAAGCTGTGCTGACAGTGTGCTGTGGATCTAG	827		
QY	840 CCTGAGGCTGATGATCAAGGCTTCTGTGACAGATGATGGGTCCGACCTGACCTGACCTCCACC	899		
DB	828 CTTGATTTTGAATGATCTTGAAGGCTTCTGTGACCAATATGATATCAACCTTGATGTGACCTCT	887		
QY	900 TGTGTCACTTTTCAATGATCATCTGCTTGAAGAGGGGTGAACCGCTGAGCCTTGGGAGCTCT	959		
DB	888 TGTGTCACTTTTGAATGATGCTGCTTCTTGAAGACGGGGAAACAGCTGTGAGCCAGGGGAACTGT	947		
QY	960 CAGCTGAACAGCTGATCTGTGAGAGGGGTGACACACCGTTGACCTTACAGTGAACAGATGAG	1019		
DB	948 CAGCTGAATGATGATCTTGAAGAGGGGTGACACACCGCTGACTTCTGTGAACAGATGAA	1007		
QY	1020 CTGGCTGTGGGCAACCGAGATGATGTTTCAAGCGGCGGACAGATGATGTCGACCTGACATCAG	1079		
DB	1008 CTGGCTGTGGGCAACCAAGAGGTGTCTAGCGCGGACAGAAATGTGTCTGATGATGCTGACGCT	1067		
QY	1080 GAGCTCCGGAATGAAGAGAGAAACACCAACCCCGGAGCGGGTGAAGCTGTGGGCAAG	1139		

1068 GAGCTCCAAAGTAGAGAAACAACACCCCGGGAACGGGTGAGCTTTCAGCAAG 1127
1140 AGGCAAGTGTCTGCAAGAGACATGCAAGGGCTGCAAGTACGCTTTCAGCCAGCCAAAG 1199
1128 AGGCAATGTCTGCAAGAGACATGCAAGGGGTTCAGATAGCTGTGAGCCAGGCAAG 1187
1200 CTGCAAGCCAGAGAGAGTTCGTCAGACCAAGCTGAGACCTTGGAGCCCGGAGAGCC 1259
1188 CTTCAGAGCCAGAGAGAGTTCGTCAGAGACAGATGAGAGAGCTTCGAGAGGAGAGCC 1247
1260 CCGCTGTGTGCTCTCTGCAAGATGACCGGCACTGACGTCGTCTTCAGAGCAGAGAGCA 1319
1248 CAGAGCGTGTGCTCTCTGCAAGATGACCGGCACTTCACCTCTCCAGC-----GAGAGCA 1301
1320 GAGGGGGGAAAGAGACACCAAGCTGAGATCTTAAAGAGCAATCTCAGAGATCTTCCG 1379
1302 GAGGGGGGAAAGAGACACCAAGCTGAGATCTTAAAGAGCAATCTCAGAGATCTTCCG 1361
1380 CCGAAGTTCCTC----- 1390
1362 CCGAAGTTCCTC----- 1390
1391 ----- 1390
1422 GAGCAGCTGTGTATACAGAGGGGCTATCCCTGGGAGAGGTAGCTGAGCTTACAGCAC 1481
1391 ----- 1390
1482 ACTGGGAGCTTCTGTGTTCCGAGAGAGCCAGAGAGAGATGATGTCTGTGATG 1541
1391 -----GAACTGTATCCGAGCTGGAA 1409
1542 TGGGATGGCAACCCCGAGCATTTTCATCATCTCACTCTTATGATTAACCTTGAACCTTGA 1601
1410 GGGGAAAGCTTCTGTATCTCTTGTCTCATGACACCACTTATGAGACACCAAGAGCC 1469
1602 GGGGATGGCTTCTCCAGCATACCTGTGTATCATCTCACTGTGTCTCTCCAGCAACCC 1661
1470 CTGACCAAGAGAGTGTGTGTCTGTCAGACAGGGCTGTGCTCCAGAGCAAGTGGTGTG 1529
1662 CTTCAGAAAGAGAGTGTGTGTCTGTCAGAGGGCTGTGCTCCAGAGCAAGTGGTGTG 1721
1530 AACCAAGAGAGCTGT 1589
1722 AACCAAGAGAGCTGT 1781
1590 AGCGGAGCGCTGTGAGAGCCGAGCAACACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1649
1782 AGTGGCGGCTTGT 1841
1650 CCACTGTGACTTCAAGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1709
1842 CCACTGTGACTTCAAGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1901
1710 CCAAGATGT 1769
1902 CCAAGATGT 1961
1770 GAGCTGT 1829
1962 GAGCTGT 2021
1830 GTGAGAGCTGT 1889
2022 GTGAGAGCTGT 2081
1890 AAGT 1949
2082 AAGT 2141
1950 CTGAGAGCTGT 2009

2142 CTGAAAGTACGAGCACTTTGGATGTCCCGAGAGAGCTGATGGATCTTATGCCGCTGC 2201
2010 GGGGGCTGTGAGCAGTATCCCGTGAAGTGAACCGGACCTGAGGCGCTTAACTAGAGCCG 2069
2202 AGCGGCTGTGAGCAGTATCCCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2261
2070 TACTCTCCGAAAGGAGCTGT 2129
2262 TACTCTCAGAGAGT 2221
2130 GGGGCTTCCCTTATCCCACTGACATATGACAGACACCGGAGATTGTGTGAGAAAGGG 2189
2322 GGGGCTTCCCTTATCCCACTGACATATGACAGACACCGGAGATTGTGTGAGAAAGGG 2381
2190 GGGGCTTCCCTTATCCCACTGACATATGACAGACACCGGAGATTGTGTGAGAGAGTGC 2249
2382 CATGT 2441
2250 TGGGCTTATGAGCTGT 2309
2442 TGGGCTTATGAGCTGT 2501
2310 ATCCGAAAGCGCATGT 2338
2502 ATCCGAAAGCGCATGT 2530

RESULT 8
ACTS140A 2715 bp ss-RNA 1linear VRL 27-APR-1993
LOCUS Fujinami sarcoma virus temperature sensitive (ts) p140 transforming
DEFINITION protein RNA, 3' end.
ACCESSION M14930.1 GI:209688
VERSION M14930.1 GI:209688
KEYWORDS transforming protein.
SOURCE Fujinami sarcoma virus
ORGANISM Fujinami sarcoma virus
REFERENCE 1 (bases 1 to 2715)
Chen, L.H., Hatada, B., Wheatley, W. and Lee, W.H.
Single amino acid substitution, from Glu1025 to Asp, of the fps
oncogenic protein causes temperature sensitivity in transformation
and kinase activity
Virology 155 (1), 106-119 (1986)
87044080
PUBMED 2877522
COMMENT Original source text: Fujinami sarcoma integrated viral DNA, clone
pFL-5 from ts FL-15.
FEATURES
source
CDS
/organism="Fujinami sarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11885"
<1..2697
/note="p140 transforming protein"
/codon_start=1
/protein_id="AA442403.1"
/db_xref="GI:209688"
/translation="PQDVTPKRVILGAPVLYLMDMAASGLRHPPOERTSTSA
AQTWLTQASERDRLPHCSAASHQDSHAKMGEPRLWCPKSHSELRLQDESLRLLE
LKKWMSQAKSDREYAGMLHMFQLEKQEGILHRAVDHSQIESWVLAQSTET
LSOTLRBAEELAAAGPLAKLILIRKQDLKRVFSQOMQLQSEYMTAQEVEKLKA
OYRSIVRDSLOAKRYKQYKQYKQYKQYKQYKQYKQYKQYKQYKQYKQYKQYKQYKQYKQYKQY
HYQKALPTLHSLKSLVLSQOEWMLVLKEIGERYCSLTSIOVEOVLATHOBAVAHEMID
PATERSSFVQCHRYSVPPAVTFDSGLLEBASLEBGLQJNELTIESVQHSUTVC
EELASRKAVSSKEQWELQVLELREBELALPGEVHLIGKRGQIGRESAQOQJVLVC
AQAQQAQKQDMTLKXLELSDPEPALPLQDRSAGCTDQERSGVTALXTKNHS
GIESPESLPPVPLPIPEVQKPLCOAWHGAIPREVOELLYKSGDFLVRSEQKQK
VLSVLMQDQRRHPTIIQAADNLVRLDEDGLRITLIDHLOSQRITRSGVILTRA
VAKDKVLANHEVULGERIGRGNPQVSGRIRADNTTPVAVVSARETLPEBKAFLQ
EARLKKQCNHNPVILIGVCTOKQPIYIMELVQGGDFLSFRSKPRLLMKKLLIOM
DNAAGMEYLESKHCHIRDLARNCLVTEKNTLKSDFQMSRQEDGVAASIGMKQI
PKWTAPALANYGWSSBSDWVSFGILLWEAFSLGAVPYANLSNOOTREAIQGVRL

Db	2172	GCAGATTCACAGAGAGCCACGAAAGCAAGAGCCGGAAAAAGCGAAGAAAAAGTACGGC	2231
Qy	592	GCAGCTGTGAAAGCTCTTTGCTGACCAACAACCGTATGTGCTGGGCTGCGGCTGGC	651
Db	2232	GCAGCTGTGAAAGCTCTATTCCTTTGCAACAACGTAAGTGTCTGGCTGTGAGGCGGCTG	2239
Qy	652	AGTCAACACCAAGCAACAACAACGACTCTGTGCTCCGCGCTGTGCTGCTCACTGACAG	711
Db	2292	CGTTGCACCAACCAACATCACTACACAGCGTGTCTGCCACCCCTGCACAGATCCCTTACA	2351
Qy	712	ACCTGCACGAGGAATGCTTGCAATCCGGAAGAGATCTCTCAGAAATCACTGGAGATTA	771
Db	2352	GCTTCGACAGAGAAATGATCTTGTGTTGAAGAGAAATCTTGAGGAGATCTGATGACATCA	2411
Qy	772	GCAGCTGTGTCAGAGATAGGTGTGTGCGCATTCACCGGAGATGAGCTTCAGACTGTGCC	831
Db	2412	CCAGTCTGTGTGACAGAGAGACGTGTGCGCATTCACAGAAAGTTGTCCACGCCCTGACGA	2471
Qy	832	GCATTCAGACCTGAGAGCTGAGTACCAAGGCTTCTGCGACAGTATGAGTCCGACCTGACG	891
Db	2472	TGATTCGACCCCGCACCTAGATACAGACAGCTTTGTTCACTGTGCACAGGTACAGCTTGAG	2531
Qy	892	TCCCAACCTGTGTCACTTTCATAGTACACTGTCTTGAAGAGGTGGAACCGCTGAGCCTG	951
Db	2532	TACACACCGGACGTACCTTTGATGAGACTTGTCTGGAAGAGCAGAGAACTTGAGCGACG	2591
Qy	952	GGAGCTTCAGACTGAAACGAGCTGACTGTGTGAGAGGCGTGCACACAACGCTGACCTCAGTGA	1011
Db	2592	GGAGCTGTGACACTGAATAGCTGTACATCGAGAGCTGTGCAACACTCTTGACATCATTTG	2651
Qy	1012	CAGATGAGCTGTGCTGTGGCCAACGAGATGATGTTTCAGGCGGACAGAGAGATGTATTGCGACG	1071
Db	2652	AGGAGAGAGCTGTGTGGCCACAGAAAGGACGATGAGAGCAAGAGACGGGTGTGTGGAGCG	2711
Qy	1072	TGCAACAGAGACTCCGGAATGAAGAAGAAACAACCAATCCCGGAGCGGCTGCAGCTGC	1131
Db	2712	TGCAAGTGTGACCTCGGGGCGAGGAGCTGCGACTAGCCCTGTGGGAGCGGCTGTGCACTTTCG	2771
Qy	1132	TGGCAAGAGGCAAGTGTGCAAGAAGCACTGCAGGGGCTGCAAGGTACCGCTGTGACGCC	1191
Db	2772	TGGCAAGAGGCAAGGAGCTGTGCGGAGGCTCCACAGCAAGCACTGAGGACTTGTGTCTGCTC	2831
Qy	1192	AGGCGCAAGCTGCAGAGCCCGACAGAGATTTCTGTGCAGAACCAACTGAGAGCACTTGAGCCCCG	1251
Db	2832	AGGCGCAAGCTGCAGAGCAAGCGGAGTATGCTGTGCCAACAACTGTGCAGAGCTGGGCTGTG	2891
Qy	1252	GCGAGCCCCCGCTGTGCTGCTCTCTGCAAGATGACCGCACTCCAAGTGTCTGTGAGAGC	1311
Db	2892	AGGAGCCCCCTCCGCTGTGCCCCCTGTGCAGAGAGACCGGCAAGTGCAGCGCTTCACAGATTC	2951
Qy	1312	AGGAGCGAGAGAGGGGGAAAGAACCAACCAAGCTGAGATCTTTAAGGCCACATCTCAGGAA	1371
Db	2952	AGGAGCGAGTGTGGGGTGA-----CTGCGCTGAAGAACATCAAGAACCAATCTTGAGGCA	3005
Qy	1372	TCTTCCGCCCCAAGTTCTCG-----	1391
Db	3006	TCTTCAAGTCCAGAGTTCTCGCTGCGCACTCCCGTCCCTCATCCCGAGAGGTGCAAGAC	3065
Qy	1392	-----	1391
Db	3066	CGCTGTGCGACAGAGCTGTGTACCAAGGGGACCAATCCCGCTCGAGAGTGCAGAGACTGC	3125
Qy	1392	-----	1391
Db	3126	TGAATAACAGCGAGACTTCTCTGTGTGCGGAGAGCCAGGGGAAAGAGATATGTCTCA	3185
Qy	1392	-----AACCTGTACC	1401
Db	3186	GCCTGTCTGTGGATGTGGGACGCCCGGACCTTCATCATCAGGCTGTGTAACAACCTGTGACA	3245
Qy	1402	GACTGGAAGGGGAAGCTTTCCTAGCAATCTTTGCTCATCGACCACTTACTGAGACCC	1461

Db	3246	GGCTGGAAGACGATGGCGCTTCCACCACTCCCGCTGCTCATGACCACTGCTGCAAGCC	3305
QY	1462	AGCAGCCCTCTACCCAGAAAGAGTGGTGTGTCTCTGCAACAGGGCTGTGCCAGACAAAGT	1521
Db	3306	AGCGGCCCATACCCCGCAGAGCGGCATCGTCTGACCAAGCGGTGCTCAAGGACAAAGT	3365
QY	1522	GGGTGCTGAAACCATGAGAGACTGGTGTGTGGGTGACAGAGATTGCACTGGGGGAACTTTTGGCG	1581
Db	3366	GGGTGCTCAACCAATGAGAGAGCGTGTCTGTGGGGGAGCCCAATTGGCGGGGGGAACTTCGGGG	3425
QY	1582	AAGTGTTCAGGGGAGCGCTGCGAGCCGACACAAACCTCGTGGCGGTGAGTACTTGTGTGAG	1641
Db	3426	AGGTGTTCAGGGGCGCCGCTGCTGCTGCTACACACCCCGTGGCGGTGAAATCTGTCCGGCG	3485
QY	1642	AGAGCTTCCCACTGACCTCAAGGCCAAAGTTTTCACAGAAAGCGAGGATCCTGAAAGAGT	1701
Db	3486	AAACCCCTGGCGCCGAGACTGAAGGCCAAAGTTTCTGACGAAAGCCAGGATCTCTCAAGCAGT	3545
QY	1702	ACAGCCACCCCAACATGTGTGGTGTCAATTGGTGTCTTGACCCCAAGACCCATCTTACA	1761
Db	3546	GCAACCAATCCCAACATGTCCTGGGTCACTCGCGCTGTCCACCCAGAAAGCACCACCAATTTACA	3605
QY	1762	TGCGATGAGGACCTGTGCAAGGGGGGCGACTTCTGACCTTCTCTCCGACGGAAGGGGGGCC	1821
Db	3606	TGCGATGAGACCTGTGCAAGGGAGGGGACTTCTTGACCTTCTTGCGACGCAAGGGGGCCCC	3665
QY	1822	GCTGTGGGGGTGAAGACTCTGTCTGCAGATGGTGTGGGGATGACAGCTGTGGGATGAGATACC	1881
Db	3666	GCTGTGAAGATGAAGAAAGCTGATCAAGATGATGGAGAAATGGCGGGCGGGGATGGAGATACC	3725
QY	1882	TGAGAGCAAGTGTGCATCCACCGGGAACCTGGTGTCTGGAACTGCTGTGTGACAGAGA	1941
Db	3726	TGGAAGCAGACATGTCATCCACAGGGACCTGGCTGCCCACTGCTGTGTGACAGAAA	3785
QY	1942	AGAAATGTCTGAATAATGATGATGACTTTTGGAGATGTCCGACAGGAAACCGGATGGGATCTATG	2001
Db	3786	AGAAACCCCTGAAATACAGCACTTTTGGAGATGTCCGCGCAGGAGAGAGATGTGTCTATG	3845
QY	2002	CAGCTTCAGGGGGGCTCAGACAAAGTCCCGGTGAAGTGAACCGCACCTGAAGGCCCTTAAT	2061
Db	3846	CCTCCACGGGGGGGATGAGCAAGATCCCCGTGAATGGAATGCGCCCCCGAAGGCTCTGAATT	3905
QY	2062	ACGGCCGCTACTCTCTCCGAAAGCGACGTGTGAGACTTTGSCATTTGCTCTGTGGAGACT	2121
Db	3906	ACGGCTGTACAGCTCGGAGAGCGACCTGTGTGAGACTTTGGAGATCTCTGTGGAGAACT	3965
QY	2122	TCACCCCTGGGGGGCTCCCCCTTATCCCAACCTCAAGCAATGACAGACACAGGGAGTTTGTGG	2181
Db	3966	TCACCCCTGGGGGGCTGGTGCCTTACCCCAACTCAAGCAACAGCAGAACCCGAGGGCCATGG	4025
QY	2182	AGAAAGGGGGGCGTGTCTGCTCCCTGCGCAAGCTGTGTCTCTGATGCGTGTTCAGAGCTCATGG	2241
Db	4026	AGCAGGGGCTGGGGCTTGAAGCCCCCGGAGACAGTGTCCCGAGAGAGTTAACCGCTGATGTC	4085
QY	2242	AGCAGTGTGGGCTTATGAGCTTGGAGCGAGCGGCCAGCTTCAAGCACCATTTACAGGAGC	2301
Db	4086	AGCGCTGTGGAGATGACACCTTACAGCGCGGCGAGACTTCTGGGGCGGTCCACACAGGAGC	4145
QY	2302	TGCAGAGATCCGAAAGCGGCAATGGTGTGAG	2331
Db	4146	TCATCGCATTCGCAACCGGACCGCTGTGAG	4175

RESULT 10				
AF033810				
LOCUS	AF033810	4788 bp	RNA	linear
DEFINITION	Fujinami sarcoma virus, complete genome.			
ACCESSION	AF033810			
VERSION	AF033810.1	GI:2801466		
KEYWORDS				
SOURCE				
ORGANISM	Fujinami sarcoma virus			
	Fujinami sarcoma virus			
	Vitruves; Retroid vitruves; Retroviridae; Alpharetrovirus.			

REFERENCE 1 (bases 1 to 4788)
AUTHORS Petropoulos, C.J.
TITLE Appendix 2: Retroviral taxonomy, protein structure, sequences, and genetic maps (in) Coffin, J.M. (Ed.); RETROVIRUSES: 757; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, NY, USA (1997)
REFERENCE 2 (bases 1 to 4788)
AUTHORS Chappey, C.
TITLE Direct Subversion Submitted (12-NOV-1997) NIH, NLM, Rockville Pike, Bethesda, MD 20894, USA
FEATURES
source . location/Qualifiers
1..4788
/organism="Pujiamt sarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11885"
22..4442
/product="p140 polyprotein"
22..102
380..3928
/codon_start=1
/product="p140 polyprotein"
/protein_id="AAC82565.1"
/db_xref="GI:2801467"
/translation="MEAVIKVTSACKTYCGKTPSKRIGAMLSQLOKRLGILMSLD
LVSFGSDPITATLTORAMVYKSGSLKTMGLVGLAKYAREQVTSBAKRMGLGSG
GRVSPGPECTEKATERRIDKGETTVQDTPKAPREIATPKVGTSGHGTALICN
CATASA.PPPYVSGSLYPSIAGVGEQGGGDTPRGAEDPRAPRGVGTGAPGALTDM
ARTREBLASTGPPVAMPVVIKTEGPAMPLEBKLTAGLAGVAGAGLSPLAVAVE
ALMSBPLPHDVTNPMRVILGPPAPHALMDAMASGLHRPOQESTSAAAGTURL
TQASBSRHLPHCSAASHODHSAMGFGELMKGTLLRLRODESLLELMKKM
SGAKSDREYAKLHMHPSQLEKQSGHBRADHSSQGESMTQASQETTLSTQLR
RAHEBLAAGPLAKLSTLRDQQLRVKVSQWQSLQSTYAMTQOEVEKLKQYRSLV
ROSTQARKYQESKDKEREKAKYRSLKLYALDNOYVLAOVAALHHHRYGRA
LPTLSLSLYLOQEMVLVKEILGEYCSITSLVOEDVLAHQVAAVAMIDPATYS
SIFQCHRDSEVPAVTPDESLEBAENIEPGLQNELTIESVHSLTISIEBELAS
RKAVSKEDRWELQVLEIRGEELASRGRVHLIGRQGLRBAQOOLQVLCVQAOLQ
AQRDMLANKLAIEGSEPPPALPQLODRSABSTDBRSGVTRKTIKHSIGISPR
FSLPPVPLIPEVQPLCCQAMVTHGAI PSEVVELKYSQDFLVRSGQGEVYSLV
WDQPRHFIIOADNLRYLEDDELPTIPLIDHLOSQRITRKSQIVTRALVLDK
VNHEDVLLGERIGRNGFSEVSGRLRABNTPAVKSCEBTLPELKAFLDEARLIK
OCHNPVIRLIGVCTOKOPYIYIMELVOGGDFLSFRSGRLKMKKLIMENNAAG
MEYLESKHCHHDLAARNCLVTEKNTLKISDFGMSRQEDGVYASTGEMKOLPVKTA
PELANYGWYSSSESDVWSFGLLMEASLSLCAVPYANISNOQTRAIISQVRLPEPQCP
EDVTRLMQRCEYDPIHRSPFGAVHODLAIIRKRR"

mat_peptide
380..835
/product="p14 MA"
899..1087
/product="p10"
1085..1309
/product="CA"
1307..3925
/product="FBS"
4196..4442

mat_peptide
38.3%; Score 1024.8; DB 14; Length 4788;
Best Local Similarity 66.3%; Pred. No. 4e-147;
Matches 1650; Conservative 0; Mismatches 612; Indels 228; Gaps 3;

QY 64 ACAGACATATGGGCTTCTTCTGAGCTGACGCCCCCAGGGCCACGGGGTCTGTAGC 123
Db 1446 ACAGTGCATGGGCTTGGGCGGAGCTGTGTGCCGAAAGGGGCAACTGAGCTGTGC 1505
QY 124 AAATGCAGAGCGGAGCTTCTCTACTGAGGGCACTGAGAAAGTGATGCCAGCGG 183
Db 1506 GGGTGCAGGACAGGAGATTGGCTCTGAGCTGAGAGAGAGATGATTCACAGGTG 1565
QY 184 TCAAGATGACAGGAGTATGACGAGCTCTTCAACCATGTCC-----CTGC 231
Db 1566 CTAAAGAGCGACCGGAGATACCGGGAGATGTGCAACCATGTCTCTCACTGGAAGAAC 1625

QY 233 AGACAGTGGGGCCAGAGCCGGGCATATGACCTTGACACGCCCATCACTAGTCTTGGG 291
Db 1626 AGGAGGGCCTGGGACATCTCCCTGCGCACGACACAGAGCCAGATGGGAGATCGTGGT 1685
QY 292 CTGAGATCACAGCCAAATGAGGGCCTGAGCCGCTTGTCTGGGACGACGAGAGATC 351
Db 1686 GGGTCTGGCAAGCCAGACTGAGACACTGAGGCAGAGCCGTGGGGCCGCGGAGAGAC 1745
QY 352 TGAAGTCAAGGGCCCTGAGCAAGCTGAGGCTGTCTATCCGGGAAAGGAGAGCTTCGGA 411
Db 1746 TGGCAGCGGGGCGGCTGGCCAGCTGAGCATACTGATCCGTGCAAGACAGACTTTCGA 1805
QY 412 AGACCTACAGAGAGAGTGGCAGAGCTGACAGAGAGCTCAACAGACCACAGCAG 471
Db 1806 AGGTCTTCAATGAGCAGTGGCAGAGCTGACGACGAGAGTAAACCTTGACACACAGCAG 1865
QY 472 ACATTGAGAGCTGAAGAGCCAGTACCGAGCTCTGGACAGGACAGTGCCTCAAGC 531
Db 1866 AAGTTGAGAGCTGAAGGCAAGTACCGAGCCTGGCTGTGACAGCACCCAGGCCAAC 1925
QY 532 GCAAGTACAGAGAGCCAGCAAGACCAAGGACGTTGACAAAGGCCAAGGACATATGTC 591
Db 1926 GCAAGTACAGAGAGCCAGCAAGACCAAGGAGCGGAAAGGCGAAGAAAGTACGTC 1985
QY 592 GCAGCTGTGAGAGCTCTTGTCTCACCAACACCGCTATGTGTGGCGCTGGCGGCTGCGC 651
Db 1986 GCAAGCTGTGAGAGCTTATGCTTGTGACACACAGTACGTCGTGTGACAGGCGGCTG 2045
QY 652 ACCTACACCAAGACACACACACAGCTCTCTGTCGCCGCTGCTGCGGTCACTGACAG 711
Db 2046 CATTGACACACACACATCACTACAGCCTGCTGCTGCCACCTGTGACGAGTCCCTTACA 2105
QY 712 ACTGACAGAGAGATGCTTGCATCTGAAAGAGATCCTGACAGAAATCCTGAGATTA 771
Db 2106 GCTTGACAGAGAGATGCTTGTGTTTGAAGAGATTTTGAGGGAGTACTGTAGATCA 2165
QY 772 GCAGCTGTGAGAGATGAGTGTGTGGCCATTACCGGAGATGAGCTGACGTGCTGCC 831
Db 2166 CCAAGTGTGAGAGAGAGAGAGCTGTGGCATCACAGAAAGTTGCCAAGCGGTGAGA 2225
QY 832 GCATTCAGCTTGAAGCTGATACCAAGGCTTCTCTGACAGATTTGGTCCGACCTGACG 891
Db 2226 TGAATCAACCCGCCCACTGATGACAGAGCTTGTGATGAGTCCACAGTTCAGTGAAG 2285
QY 892 TCCCAACCTGTGACGTTGATGAGTCACTGTGAGAGAGGTGAACCGCTGAGAGCTG 951
Db 2286 TACCAACCGGACAGTGAATTTGATGAGAGCTTGTGAGAGAGGACAGAACTGAGAGCAG 2345
QY 952 GGGAGCTCCAGCTGACAGAGCTGATGTGAGAGAGCTGACGACAGCTGACCTCAATGA 1011
Db 2346 GGGAGCTGACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2405
QY 1012 CAGATGAGCTGTGTGGCCACAGAGATGTGTTACAGGGGGCAGAGATGTGTTACAGCAG 1071
Db 2406 AGAGAGAGCTGTGTGGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2465
QY 1072 TGCACAGAGAGCTCGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1131
Db 2466 TGCAGAGTGAAGCTGCGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2525
QY 1132 TGGGCAAGAGAGAGAGTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1191
Db 2526 TGGGCAAG 2585
QY 1192 AGGCAAGTGCAGAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1251
Db 2586 AGGCAAGTGCAG 2645
QY 1252 GCGAGCGCCGCTGTGTGCTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1311
Db 2646 AGGAGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2705

QY		1312	GGAGCGGAGAGGGGGGAAAGACACCAGCTGGAGATCTCTTAAGAACAATCTTCAGAA	1371
Db		2706	AGGAGCCGAGTGGGGGTGA-----CTGGCTGAAGACCATCAAGAACCAATCTCGGCA	2759
QY		1372	TCCTCCGCCCCAAAGTCTCG-----	1391
Db		2760	TCTTCAGTCCAGGTTCTCGCTGCACCTCCCGTGCCCTCATCCCGAAGGTGCAGAAAGC	2819
QY		1392	-----	1391
Db		2820	CGCTGTGCCAGCAGGCTGTATACACGAGGGCCATCCGCGCTCGGAGGTGACGAGACTGC	2879
QY		1392	-----	1391
Db		2880	TGAATAACAGCGAGACTTCTCTGTGCGGAGAGCACAGGGAGAACAGATATGTCTCA	2939
QY		1392	-----AACCTGATCC	1401
Db		2940	GCGTGTGTGGGATGGGCAAGCCCCGGACATTATCATTCAGGCTGTGACAACTTGATCA	2999
QY		1402	GACTGGAAGGGGAAAGGCTTTTCTTAGCATTTCTTTGCTATCGAACCACTTATCGACACCC	1461
Db		3000	GCGTGAAGAAGATAGGCTTCCACATCCGCGTGTATGACCACTGTGTGCAGAGCC	3059
QY		1462	AGCAGCCCCCTCACCAAGAAAGTGTGTGTCTCTGCACAGGGCTGTGCCACAGACAAGT	1521
Db		3060	AGCGGCCCATTACCCGCAAGAGCGGCATCTCTGACCAAGAGCCGTGTCTCAAGACAAGT	3119
QY		1522	GGGTGTGGAACCATAGAGGACCTGGGTGTGGGTGAGACAGATTGGAACGGGGGAACTTTGGCG	1581
Db		3120	GGGTGTCTCAACATAGAGGACGTGTCTGTGGGGAGCGCATTTGGCCGGGGGAACTTCGGGG	3179
QY		1582	AAGTGTTCAGCGGACGACTGTGAGCGGCAACAACCTGTGTGGCGGTGAAGTCTTGTGAG	1641
Db		3180	AGGTGTTCAGCGGCGCGCTGTGTGTGACAAACCCCGTGTGGGTGAATCTGTCCGGG	3239
QY		1642	AGACGCTCCCACTGACCTCAAAGCCMAAGTTCTTACAGAAAGCAAGATCTTGAACAAGT	1701
Db		3240	AAAACCTGTCCGCGAGCTGAAGGCCAAAGTTCCTGCAGAAAGCAGAGATCTTCAAGCAGT	3299
QY		1702	ACAGCCACCCCAACATGTGTGTGTCTCATTTGGTGTCTGACCCCAAGAGGCCCATTTACA	1761
Db		3300	GCAACCAATCCCAACATGTGTGTGTGTCTGTGTGTGACCCCAAGAGAGGCCCATTTACA	3359
QY		1762	TGCTCATGGAAGCTTGTGAGGGGGGCGACTTCTTGAACCTTCTCTCGCACGAGAGGGGCCC	1821
Db		3360	TGCTCATGGAAGCTGTGTGAGGAGGGGACTTCTTGAAGCTTCTGTGTGTGACAGAGGGGCCC	3419
QY		1822	GCTGTGCGGGTGAAGACTTGTGTGACAGATGTGTGGGGATGTACAGCTGTGGAATGAGTACC	1881
Db		3420	GCTGTGAAGATGAAGAACTGTATCAAGATGATGGAAGATGCGCGGGCGGACATGAGTACC	3479
QY		1882	TGGAAGAGCAAGTGTGTGATCCAACGGGACCTGTGTGTGTGCGGAATGTGCTGTGTGACAGAGA	1941
Db		3480	TGGAAGAGCAAGTGTGTGATCCAACGGGACCTGTGTGTGTGCGGCAACTGTGTGTGACAGAAA	3539
QY		1942	AGAAATGTCTGAAGTTCAGTACCTTTGGGAGTGTCCGAGAGAGAAAGCGAATGGGGCTTATG	2001
Db		3540	AGAAACCTCTGAAGTTCAGGCACTTTGGGAGTGTGTGCGGACAGAGAGAGATGTGTGTATG	3599
QY		2002	CAGCCTCAGGGGGGCTCAGACAAGTCCCGCTGAAGTGAACCGCACTTGAGGCCCTTAACT	2061
Db		3600	CCTCAACGGGGGGGATGAAGACAGATCCCGTGAATGGAATGTGCCCGGAGGCTGTGAATT	3659
QY		2062	ACGGCCGCTACTCTCTCGAAGCGACGTGTGAAGCTTTGGCATTTGTCTGTGGAGACTT	2121
Db		3660	ACGGCTGTGTAGCTCTCGGAGAGCGACGTGTGAAGCTTTGGGATCTGTGTGTGGAAAGCTT	3719
QY		2122	TCAGGCTTGGGGGCTCCCCCTATCCCAACCTCAGCAATCAGACAGACAGGAGTTTGTGG	2181
Db		3720	TCAGGCTTGGGGGCTGTGCTTACGCAACCTCAGCAACACAGACAGCGCGAGGCAATCG	3779
QY		2182	AGAAAGGGGGGCGGTGTGCTGTGCCCCAGAGCTGTGTCTGATGTCCGTGTTCAGGCTCATGSG	2241

Db	3780	AGCAGGGCGTGCAGCTGAGAGCCCGGAGGAGCTGCCGAGAGAGCTGACCGCTGATGC	3839
Qy	2242	AGCAGTGTGGGCGCTATGAGCGCTGGGAGCGGCGCCAGCTTCAGCAGCATTACAGGAGCG	2301
Db	3840	AGCGCTGTGGGAATTAAGACCTTCACAGGGGGCGGAGCTTCGGCGCGGTCCAGCAGGAGCC	3899
Qy	2302	TGCAGAGCATTCCGAAAGCGGCGCATGGTGAG	2331
Db	3900	TCATCGGCATCCGCAACGGCAGCGCTTGAG	3929
RESULT 11			
LOCUS	A06935	856 bp	linear
DEFINITION	fee/fps-related mRNA.		
ACCESSION	A06935		
VERSION	A06935.1		
KEYWORDS	GI:490473		
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	other sequences; artificial sequences.		
AUTHORS	1 (bases 1 to 856)		
TITLE	Van de Ven, W.J.M., Roebroek, A.J.M. and Schalken, J.A.		
JOURNAL	Recombinant DNA and cDNA, mRNA, protein, antibodies, and a method		
FEATURES	Patent: BP 0246709-A 2 25-NOV-1987;		
source	Stichting Katholieke Universiteit		
	Location/Qualifiers		
	1..856		
	/organism="synthetic construct"		
	/mol_type="unassigned RNA"		
	/db_xref="taxon:32630"		
	3..638		
	/note="Protein sequence is in conflict with the conceptual		
	translation"		
	/codon_start=1		
	/transl_table=1		
	/product="fee/fps-related protein"		
	/protein_id="CAA00604.1"		
	/db_xref="GI:4526613"		
	/translation="KQYSHNYRLIVCTQKQPIYIVMELVGGDFLTLRTRGARL		
	/translatability="KQYSHNYRLIVCTQKQPIYIVMELVGGDFLTLRTRGARL		
	RVKTLQLQVGDAAAGMEYLSKCCIRHDLAARCLVTERKVLKISDGMRENDGVY		
	AAAGGLQVDPVAKTAPALNYGRYSFSDVMSFGILMEFSLGASPYPMISNOTRE		
	FVEKGRGLPCELCPDAVFLRMQCMWYEPGRPSFSTIYQELQSIKRRH"		
ORIGIN			
Query Match	32.04; Score 856; DB 6; Length 856;		
Best Local Similarity	100.0%; Pred. No. 2.5e-121;		
Matches	856; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1693	TGAAGCAGTACAGCAGCCCAACATCGTGGTCCATTGTTGTTGTGACCCAGAGAGAGC	1752
Db	1	TGAAGCAGTAAAGCCACCAACATCGTGGTCCATTGTTGTTGTGACCCAGAGAGAGC	60
Qy	1753	CCATCTACATCGTATGAGAGCTTGTGAGGAGGAGGAGACTTCCTGACCGACGG	1812
Db	61	CCATCTACATCGTATGAGAGCTTGTGAGGAGGAGGAGACTTCCTGACCGACGG	120
Qy	1813	AGGGGGCCCCCGCTGCGGGGTGAAGCTTGCTGTCAGATGTGTGGGGATGACAGCTGCGCA	1872
Db	121	AGGGGGCCCCCGCTGCGGGGTGAAGCTTGCTGTCAGATGTGTGGGGATGACAGCTGCGCA	180
Qy	1873	TGAGAGTACCTGGAGAGAGAGAGCTGTCATCCACCGGGAGCTGGCTGCGAACTGGCTGG	1932
Db	181	TGAGAGTACCTGGAGAGAGAGAGCTGTCATCCACCGGGAGCTGGCTGCGAACTGGCTGG	240
Qy	1933	TGACAGAGAAAGATGTCTCTGAAGATCAGTACTTTGGAGATGTCCCGAGAGAAAGCGATG	1992
Db	241	TGACAGAGAAAGATGTCTCTGAAGATCAGTACTTTGGAGATGTCCCGAGAGAAAGCGATG	300
Qy	1993	GGGTCTATGACGCTCAGGGGGGCTCAGACAGTCCCGGTGAAGTGAACGCGACCTGAGG	2052

Db 301 GGGCTATGACAGCTCAGGGGGGCTCAGACAAAGTCCCCGTGAAGTGAGCCGACCTGAGG 360
Qy 2053 CCGTTAACTAACGGCCCTACTCTCCGAAAGCGACGTGAGACTTTGGCATTTTGGCTCT 2112
Db 361 CCGTTAACTAACGGCCCTACTCTCCGAAAGCGACGTGAGACTTTGGCATTTTGGCTCT 420
Qy 2113 GGGAGACCTTCAAGCTGAGGGGCTCCCTCATCCAACTCAGCAATCAGACAGACGGG 2172
Db 421 GGGAGACCTTCAAGCTGAGGGGCTCCCTCATCCAACTCAGCAATCAGACAGACGGG 480
Qy 2173 AGTTTGTGAGAAAGGGGGGCGGTCTGCTCCCAAGCTGTCTCTGATCCGTGTCA 2232
Db 481 AGTTTGTGAGAAAGGGGGGCGGTCTGCTCCCAAGCTGTCTCTGATCCGTGTCA 540
Qy 2233 GGGCTATGAGAGAGTCTGGGCTATGAGCCTGAGCGGCGCCAGCTTCAGACCATCT 2292
Db 541 GGGCTATGAGAGAGTCTGGGCTATGAGCCTGAGCGGCGCCAGCTTCAGACCATCT 600
Qy 2293 ACCAGAGCTGACAGACATCCGAAAGCGGATCGTGAGGCTGGACCCCTTCTCAAGC 2352
Db 601 ACCAGAGCTGACAGACATCCGAAAGCGGATCGTGAGGCTGGACCCCTTCTCAAGC 660
Qy 2353 TGGTGCCTCTGACAGGCTTGAAGTGAGCTCTCAGCGGCTCCAGCTCATATGCTGACAGC 2412
Db 661 TGGTGCCTCTGACAGGCTTGAAGTGAGCTCTCAGCGGCTCCAGCTCATATGCTGACAGC 720
Qy 2413 TCTTCAAGTCCAGACCTCCTGACACAGCATCAGCTGCGGCGAGAGTGACGGCGCT 2472
Db 721 TCTTCAAGTCCAGACCTCCTGACACAGCATCAGCTGCGGCGAGAGTGACGGCGCT 780
Qy 2473 GTCCTCTGTGTGCTCCCTGCTGCTGCGAGGCTTCTTCCGGGCGAGAAACAATAAAC 2532
Db 781 GTCCTCTGTGTGCTCCCTGCTGCTGCGAGGCTTCTTCCGGGCGAGAAACAATAAAC 840
Qy 2533 ACTTGTGCCCACTGAA 2548
Db 841 ACTTGTGCCCACTGAA 856

RESULT 12
M22820
LOCUS M22820 4901 bp DNA linear UNA 04-AUG-1993
DEFINITION Figure 4. Nucleotide sequence of the proviral DNA of the variant clone of ST-FesV.
ACCESSION M22820
VERSION M22820.1 GI:340749
KEYWORDS
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 4901)
AUTHORS Van Groningen,J.J., Van den Ouweland,A.M., Verbeek,J.S., van der Kemp,A.W., Bloemeste,H.P. and Van de Ven,W.J.
TITLE Structural analysis of a variant clone of Snyder-Theilen feline sarcoma virus
JOURNAL Virus Res. 8 (4), 349-361 (1987)
MEDLINE 88129048
PUBMED 2829460
FEATURES
source Location/Qualifiers
1..4901
/organism="unclassified"
/mol_type="unassigned DNA"
/db_xref="taxon:12908"

ORIGIN
Query Match 32.0%; Score 854.4; DB 13; Length 4901;
Best Local Similarity 77.7%; Pred. No. 4,7e-121;
Matches 1159; Conservative 0; Mismatches 121; Indels 212; Gaps 2;

Qy 1104 ACCCAACCCCGGAGCGGGTGCAGCTCTGGCGAAGAGGCAAGTGTCTGCAAGACACTG 1163
Db 1874 ACCCAACCCCGGAGCGGGTGCAGCTCTGGCGAAGAGGCAAGTGTGTGCAAGAGCGGCTG 1933

Qy 1164 CAGGGGCTGACAGTAGCGTGTGACGCCAGGCCAAGCTGCAGGCCAGAGAGTTGCTG 1223
Db 1934 CAGGGGCTGACAGTAGCGTGTGTGTGACGCCAGGCCAAGCTGCAGGCCAGAGAGTTGCTG 1993
Qy 1224 CAGACCAAGCTGAGACCTGGGCCCCGGCGAGGCCCTGTGTCTGCTCTGCGAGAT 1283
Db 1994 CAGGCGCAAGCTGAGACCTGGGCCCCGGCGAGGCCCTGTGTCTGCTCTGCGAGAT 2053
Qy 1284 GACCGCACTCAGAGTGTCTCTGAGCGAGAGCGAGAGAGGGGGAAAGACACCAAGCTG 1343
Db 2054 GACCGCACTCAGAGTGTCTCTGAGCGAGAGCGAGAGAGGGGGAAAGACACCAAGCTG 2113
Qy 1344 GAGATCTTTAAGAGCCACATCTCAGGAATCTTCGGCCCAAGTCTCG----- 1391
Db 2114 GAGATCTTTAAGAGCCACATCTCAGGAATCTTCGGCCCAAGTCTCGCTCCCTCCAGCC 2173
Qy 1392 ----- 1391
Db 2174 CTGACAGCTGTACAGAGGTGAGAAAGCCCTGACAGACAGCTGTGTAAACAGGGGCGC 2233
Qy 1392 ----- 1391
Db 2234 CTCCACGGGCGAGAGGTGAGCTGAGCTGTGAGCACTCTGGGAGCTTCTGATGGCGAG 2293
Qy 1392 ----- 1391
Db 2294 AGCAAGGCGAGAGGAATATGTGCTGTGTGTGGAGCGGCGAGCCCGCACTTC 2353
Qy 1392 -----AACTGTACCGACTGAGAGGGGAAAGGCTTCTTACATTCCT 1433
Db 2354 ATCATTCAGTCCGCTGACAACTCTTACCGACTGAGAGGGGAGGTGCTTTCAGACATCCC 2413
Qy 1434 TTGCTCATGACCAACCTACTGAGACCCAGAGCCCTTACCAAGAGAGAGTGTGTTC 1493
Db 2414 TTGCTCATGACCAACCTACTGAGACCCAGAGCCCTTACCAAGAGAGAGTGTGTTC 2473
Qy 1494 CTGACAGGCTGTGTGCCAAGACAAGTGTGTCTGAACCATAGAGACCTGTGTGGGT 1553
Db 2474 CTCAACAGGCTGTGTGCCAAGACAAGTGTGTCTGAACCATAGAGACCTGTGTGGGT 2533
Qy 1554 GAGCAGATTGACAGGGGGAACCTTTGGCGAAGTGTAGAGGAGCGCTGCGAGCCGCAAC 1613
Db 2534 GAGCAGATTGAGCGGGGGAACCTTTGGCGAAGTGTAGAGGAGCGCTGAGGGCGCAAC 2593
Qy 1614 ACCCTGTGGCGGTGAAGTCTTGTGAGAGAGCGCTCCAGCTGACCTCAAGGCCAAGTTT 1673
Db 2594 ACTTATGTGCGGTGAATCTTGTGCGAGACACTTCCAGCTGACATCAAGGCCAAGTTT 2653
Qy 1674 CTACAGAAAGCAGAGATCTTGAAGCAGTACAGCAACCCCAACATCGTGCCTCATTTGT 1733
Db 2654 CTTCAGAGAAAGCAAGATCTTGAAGCAGTACAGCAACCCCAACATCGTGCCTCATTTGT 2713
Qy 1734 GTCTGACCCAGAAAGAGCCCATCTTACATGCTCATGTAGAGCTTGTGAGGGGGGCACTTC 1793
Db 2714 GTCTGACCCAGAAAGAGCCCATCTTACATGCTCATGTAGAGCTTGTGAGGGGGGCACTTC 2773
Qy 1794 CTGACCTTCTCTCGAGCGAGAGGGGGCGGCTGCGGGGTGAAGCTGTGCTGACATGTGTG 1853
Db 2774 CTGACCTTCTCTCGAGCGAGAGGGGGCGGCTGCGGGGTGAAGCTGTGCTGACATGTGTG 2833
Qy 1854 GGGGATGACAGCTGCTGAGTGAAGTACTCTGAGAGAGCAAGTCTGATCCAGCGGAGCTTG 1913
Db 2834 GGGGATGAGCGGCGCGGAGTGAAGTACTCTGAGAGAGCAAGTCTGATCCAGCGGAGCTTG 2893
Qy 1914 GCTGCTCGGAATCTGCTGTGTGACAGAGAAATGTCTTGAAGTCAAGTCACTTTGGGATG 1973
Db 2894 GCTGCTCGGAATCTGCTGTGTGACAGAGAAATGTCTTGAAGTCAAGTCACTTTGGGATG 2953
Qy 1974 TCCCGAGAGAAAGCGGATGGGCTTATGAGCTTCAAGGGGGCTTGAACAAGTCCCGGTG 2033
Db 2954 TCCCGAGAGAAAGCGGATGGGATCTTACGCGGCTTCAAGGGGGCTTGAACAAGTCCCGGTG 3013
Qy 2034 AAGTGACCGCACTGAGCCCTTAACTAGCGGCGCTACTCTCCGAAAGCGAGCTGTGG 2093

DB 2158 ACCTGTGCGCCCTGTAATCTTCTCCCGAGACACTCCCACTGACATCAAGGCCAAGTTT 2217
 QY 1674 CTACAGAAAGCAGAGATCTTGAAGCAGTACAGCCACCCCAACATGCTGCTCATTTGTT 1733
 DB 2218 CTTACAGAAAGCAAGATCTCGAAGCAGTACAGCCACCCCAACATGCTGCTCATTTGTT 2277
 QY 1734 GTCCTGACCCAGAAAGAGCCCATCTACATCTGATGAGAGTTGTGAGAGGGGGCGACTTC 1793
 DB 2278 GTCCTGACCCAGAAAGAGCCCATCTACATCTGATGAGAGTTGTGAGAGGGGGCGACTTC 2337
 QY 1794 CTGACCTCTCTCCGACGAGAGGGGGCCGCTGCGGGTGAAGACTGTGCTGACATGTGTG 1853
 DB 2338 CTGACCTCTCTCGTAGAGAGAGAGAGAGCCCGCTGCGGATGAAGAGCGTCTGCGAGATGTG 2397
 QY 1854 GGGGATGACAGCTGCTGCGATGAGTACCTGAGAGAGCAAGTGTGATCCACCGGAGCTTG 1913
 DB 2398 GGGGATGACAGCTGCTGCGATGAGTACCTGAGAGAGCAAGTGTGATCCACCGGAGCTTG 2457
 QY 1914 GCTGCTCGGAACTGCTGCTGAGAGAGAAATGTCTGAAGATCACTGATCTTTGGGATG 1973
 DB 2458 GCTGCTCGGAACTGCTGCTGAGAGAGAAATGTCTGAAGATCACTGATCTTTGGGATG 2517
 QY 1974 TCCCGAGAGAGAGCCGATGGGGGTCTATGACGCTCAGAGGGGGCTCAGACAAAGTCCCGTG 2033
 DB 2518 TCCCGAGAGAGAGCCGATGGGGGTCTATGACGCGGCTCAGAGGGGGCTCAGACAAAGTCCCGTG 2577
 QY 2034 AAGTGAACGCGACCTGAGGCGCTTAACTACGCGCGCTACTCTCCGAAAGCAGCTGTGG 2093
 DB 2578 AAGTGAACGCGACCTGAGGCGCTTAACTACGCGCGCTACTCTCCGAAAGCAGCTGTGG 2637
 QY 2094 AGCTTTGGCATCTTGTCTTGGAGACCTTACGCTTGGGGGCTTCCCTATCCCAACTTC 2153
 DB 2638 AGCTTTGGCATCTTGTCTTGGAGACCTTACGCTTGGGGGCTTCCCTATCCCAACTTC 2697
 QY 2154 AGCAATCAGAGACAGGAGGTTTGGAGAGAGGGGGCGCTGCTGCGCTGCCAGAGCTG 2213
 DB 2698 AGCAATCAGAGACAGGAGGTTTGGAGAGAGGGGGCGCTGCTGCGCTGCCAGAGCTG 2757
 QY 2214 TGTCTGATGCTGCTGCTTACAGCTCATGAGACATGCTGGGCTATGAGCTTGGGACAGCG 2273
 DB 2758 TGTCTGATGCTGCTGCTTACAGCTCATGAGACATGCTGGGCTATGAGCTTGGGACAGCG 2817
 QY 2274 CCGAGCTTACAGCACTTACCGAGAGCTGCAAGACATCCGAAAGGGGATCGGTGAGGC 2333
 DB 2818 CCGAGCTTACAGCACTTACCGAGAGCTGCAAGACATCCGAAAGGGGATCGGTGAGGC 2877
 QY 2334 TGGGACCCCTTGTCAAGCTGATGAGCTTGCAGAGCCCTAGGTGAGCTGCTC 2385
 DB 2878 -CGGCTCTGCTTCAAGCCGCGGCTCGACAGGCAAGATCGTACTCTC 2928

RESULT 15
 LOCUS BC073445 4788 bp mRNA linear VRT 03-AUG-2004
 DEFINITION Xenopus laevis MGC80946 protein, mRNA (cDNA clone MGC:80946
 IMAGE:5516233), complete cds.
 ACCESSION BC073445
 VERSION BC073445.1 GI:49116710
 KEYWORDS MGC.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopus;
 Xenopus laevis; Xenopus.
 1 (bases 1 to 4788)
 Kleinf, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
 and Richardson, P.
 Genetic and genomic tools for Xenopus research: The NIH Xenopus
 Initiative
 JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
 PUBMED 12454917
 REFERENCE 2 (bases 1 to 4788)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klauener, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schlier, G.D.,
 Altschul, S.F., Zeeberg, B., Bucow, K.H., Scheffer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, O., Hsieh, F.,
 Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Ueda, T.B., Toshiyuki, S.,
 Carinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mulhany, S.J., Bosak, S.A., McKean, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Sanchez, A., Whiting, M., Madan, A., Touchman, J.W., Green, E.D.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalske, U., Smallus, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932

JOURNAL
 PUBMED
 AUTHORS
 TITLE
 JOURNALREMARK
 COMMENT

3 (bases 1 to 4788)
 Klein, S. and Gerhardt, D.S.
 Direct Submission
 Submitted (14-JUN-2004) National Institutes of Health, Xenopus Gene
 Collection (XGC), National Institute of Child Health and Human
 Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
 20892-7510, USA
 NIH-MGC Project
 Contact: XGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Igor David
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@gscc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
 Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
 Featherstone, Malachi Griffith, Orl Griffith, Ran Guin, Nancy Liao,
 Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
 Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu,
 Parvaneh Seedi, JR Santos, Angelique Scherch, Ursula Skalske,
 Duane Smallus, Jeff Stott, Miranda Tsai, George Yang, Jacques
 Schein, Asim Siddiqui, Rob Holt, Marco Marra.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 155 Row: J Column: 3
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, similarity but not identity to protein.
 Location/Qualifiers
 1..4788
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="MGC:80946 IMAGE:5516233"
 /issue_type="Embryo, stage 31/32, Xenopus"
 /clone_lib="NICHG_XGC_Embd"
 /lab_host="PHIOB"
 /note="Vector: pCMV-SPORT6"
 1..4788
 /gene="MGC80946"
 /db_xref="LOCUSID:444292"
 399..2867
 /gene="MGC80946"
 /codon_start=1
 /product="MGC80946 protein"
 /protein_id="AAH73445.1"
 /db_xref="GI:49116711"
 /db_xref="LOCUSID:444292"

```

/translation="MGFAELQCPNGHATLRLDQDELHVMESMKCFIQRKDKKEY
SMNHQISVQVEKLDQSLTPGSEVSSQSDSESMATLVSTQILRSGSEVAVNAP
LKTLLIREKQOLKKSYSSEOMQMODMKTTOOIEFLRCOYSORVETROSKRY
QKACRDKDDKAKKRYKATWKLHNLHNYLAVRQJHHEHPIYSPLTHESMOS
IHQEWQILKNTMOEISEITSLQVEEVVLAHRELIASAVRISPEQSYDPLFKSIVS
ELPPIVTPFSLLEBTDLEBELQMLDVLVESVQHTLVSEBELSVTAIISKODIV
KOLHSDIAIEKRTDCREQLNLAKRHALDLYQHTMLCKRKNLOGRMLKRMN
TGFQDPALLODODNOSITSVERECGKLENIKNHLSGIFRPFSLPLPLVPD
VQKPLSQNMWHGAIIPRSFVQGLVNRGQVIRESGQKQVYLSVIMDQPPHFIION
VNLIRLBEGBSTPIPLINHPYKTOAATYKSGVINKAIYKDKRMHEHEVILGER
IGKNGNVEFSQRLADNTPVAVKSCRDLPDLKQKFLMEARILQVISHPNVKLIG
VCTQKHPILYIMELVQGDQFQTLQNEGRLKVELIRVSENAAMELESKHCHIR
DLAARNLITERKALISDFQMSREBEDVYSTGEMQOIPKMTPEALANTQVSSSE
SDVMSFGLIMEAFSLGSPVYAMATQOREAIEQGVRLVDPNCPDEYVSLMRQWE
YPKPKRPNRSIYHQVLTIRKRYR"

```

ORIGIN

Query Match 24.7%; Score 659.6; DB 5; Length 4788;
 Best Local Similarity 57.6%; Pred. No. 3e-91;
 Matches 1426; Conservative 0; Mismatches 834; Indels 216; Gaps 5;

```

QY 72 ATGGGCTTCTCTTCTGAGCTGTGACGCCCCGAGGGCCAGGGGTCTTGCAAGCAATGCAAG 131
DB 399 ATGGGGTTTGGCGAAGATTACAGTCCCAATGGGCAATGCTATTGTTACGTTTACAA 458
QY 132 GAGGCCGAGCTTGCTCACTGAGGGGATGAGAAAGTGATGGCCGAGGGGTCAAGAGT 191
DB 459 GACACTGAACCTTAGATGATGAGTCAATGAAATGTTTATACAGCTGCAAAAGGCG 518
QY 192 GACAGGAGATATGACAGAGTCTTCAACAATGTCCCTGCAGG----ACAGTGGGGCCA 247
DB 519 GACAAAGAAATATCGAATATGCTTCAACAATTTCTGTGAGGTGCAAAAAGCTGATCAG 578
QY 248 GAGCCGGGCATAGCCCTGAC-----AGCCCATAGTCAATGTCCTGGGTGAGATC 299
DB 579 TGCTTACCCCTGATATCTGAATACTCAAGCAGCTCAGTGAATGTCGGGCGCAAGTTG 638
QY 300 ACCAGCCAACTGAGGGCCTGAGCCGCTTGCTGCGGAGACGACAGAGATCTGAATCTCA 359
DB 639 GTGTCTCAGACAGAACTTACAGCATTTTACGAAAGCATTCAGAGATGTAATGCA 698
QY 360 GGGCCCTTGAAGCAAGTGAAGCTGCTCATCCGGGAAACGGCAGAGCTTGCAAGACTTAC 419
DB 699 GGAACCCCTGAGCAAAATTAACATTCTTAATCGGAAAGAGCAGGCTAAAGAAAGTCTTAT 758
QY 420 AAGGACACATGAGGAGCAAGCTGCAAGAGCTCAACAAGCCACAGCCAGCATTTGAG 479
DB 759 AGTGACCAATGGCACTTCTCAACAGGATTAATGAAAGCAACACAGACGATATCGAA 818
QY 480 AAGCTGAAGAGCCAGTACAGAGCTTGGGACGGGACAGTGCCCAAGCCAAAGCCAAAGTAC 539
DB 819 AAGCTAAGGTGTAGATGCGACCCAAATGAAAGAACTTTCCAAATCAAGAAAGAAATAC 878
QY 540 CAGGAGGCCAGCAAGAAACAAGACCGGTGACAGAGCCAAAGATATGTGCGAGCTTG 599
DB 879 CAAGAGCATATGACAGATTAAGGACCGGATTAAGCAAGAGAAATATGAAAGAAATACA 938
QY 600 TGAAGCTCTTGTCTCAACCAACCGGTATGTGCTGGGCGTGGGGCTGCGACGCTACAC 659
DB 939 TGAAGATGATGCTACATGCAATCACTATGTCTGCTGTAGAACTGCTCAGTTACAC 998
QY 660 CACGACACACACACAGAGCTCTGCGCGGCGCTGTGCGGTCACTGACGAGACCTGAC 719
DB 999 CATGAGACCACTTTTCTGTATAGTGTGCAACACTGACGAGTCCATGACAGACATTCAC 1058
QY 720 GAGGAGATGGCTTGTGCTGAAGAGAGATCTGACAGAAATACCTGAGATTAAGCAGCTG 779
DB 1059 CAGGAGATGGTCCAGATTTCTCAAGAAATATATGCAAGAAATACGTGAAATCAACAATCTA 1118
QY 780 GTGACGAGATGAGGTGGTGGCCATTCACCGGAGATGGCTGACGTGCTGCGCCGATCCAG 839
DB 1119 GTACAGAGAGAAATAGTTTCCCTGCAATGAGAAATATGCGCAGTGTCTTAACAGAAATCAGC 1178

```

```

QY 840 CCTGAGGCTGAGTACCAAGCTTCTCTGCGACAGTATGGGTCCGACCTGACGTCCACCC 899
DB 1179 CTTGAGGAGGAATATAGCAATTTTCTCCAGATGAAAGGCTAGTGTCTGAATTTGCCACC 1238
QY 900 TGTGTCAAGTTGATGAGTCACTGTCTTGAGAGGGGTGAACCGCTGAGAGCTGGGAGCTC 959
DB 1239 ATTGTGACATTTTGAATACAGATTTGTGCGGAGGAGACAGATGAGACTGGA---AAGAAACTTA 1295
QY 960 CAGCTGAACGAGCTGACTGTGAGAGCGGCGACGACAGCTGACCTCACTGACAGATGAG 1019
DB 1296 CAGCTTAATGACTTACAGCTTGAAGTGTGCAACATACCTTGACTTGTGAAGAGAA 1355
QY 1020 CTGAGCTGTGGCCACCGAGATGGTGTTCAGGCGCGACGAGATGTTACGCACTGCAACAG 1079
DB 1356 CTTGCTTGATGACAGCGGCGCATTTGAAAGCAACAGGATATATGTGAACAACTACACTCC 1415
QY 1080 GAGCTCCGGAATGAAGAGAGAAACACCCACCCCGGAGCGGTGCACTGCTGGCGAAG 1139
DB 1416 GATATTTGCAATGAGAAAGAAAGACACAGATTTGACAGAGAACAGCTTAATCTGTTGGCCAA 1475
QY 1140 AGGCAAGTGTGCAAGAAAGCACTGACAGGGGCTGCAAGTACGCTGTGACGCGCGCAAG 1199
DB 1476 AACAATGCTTACGGGATGCTCTTGTGAGATACATGATCTCTTGCAACAAAGAAAGAG 1535
QY 1200 CTGACAGGCCAGCAGAGATTTGCTGACAGACCAAGCTGAGACACTGGGCCCGCGAGGCC 1259
DB 1536 ATCAAGGGCCAGAGAGAAATGCTTAAGAGAAAGATGAACATATTGGATTGGGGACCT 1595
QY 1260 CCGCTGTGTCTGCTCTGCGAGATGACCGCCACTCACTGCTCTCGGAGCAGAGAGCA 1319
DB 1596 CCTCCTGCCCCCACTTCAAGATGATTAACAGTAACTTCACTTCACTGAGGCGCAATGT 1655
QY 1320 GAGGGGGGAGAGACACCCACGCTGAGAA----- 1347
DB 1656 GGAAGAAACATTTGAAATATTAAGAAATCACTTTCTGGAATATTTGGCGCTAAATTT 1715
QY 1348 -----TCTTAAGAGCCACATCTCAGAAATCTTCCGCCCAAGTTCTC 1390
DB 1716 TCTTTGCTCCTCTTTTACCTTTGATGTCAGATGTCAGAAACCTTTCAGCCAAAGAAAG 1775
QY 1391 G----- 1391
DB 1776 TGGTACATGATGCAATCTGCTCTGAAGTCAAGAGACTGTGGTGAATGAGGGGAT 1835
QY 1392 ----- 1391
DB 1836 TTCTTTTCCGGAGAGCCAGGGCAAGCAAGATATGTCTGTCTGTATATGAGATGGG 1895
QY 1392 -----AACTGTACGACTGGAAGGGGAGGC 1418
DB 1896 CAACCAAGCATTTTATATCTAGAACGTAATTAATCTTTACAGGTTAAGAGAGAAAGGC 1955
QY 1419 TTTCTGAGATTCCTTTGTCTATCGAACACTTACTGAGACCCAGAGGCCCTCACAAAG 1478
DB 1956 TTTTCAACGATCCCATTTGCTGATCAATCACTTTGTCAAAACCCAGCAGACAGTGAACAAA 2015
QY 1479 AAGAGTGTGTGTCTGACAGGGGTGCGCCAAAGGACAGAGTGGGTGCTGAACCATGAG 1538
DB 2016 AAGAGTGTGTATCAATAATTAAGCTATTTGTAAGAACAAATGGGTGATGGACATGAA 2075
QY 1539 GACTGTGTGTGGGTGAGCAGATTTGACCGGGGAACTTTGGCGAAGTGTTCAGCGGACGC 1598
DB 2076 GAAATCATATCTGGAGAAAGAAATTTGGAAGGAAATTTTGGTGAAGTATTCAGCGGAGCC 2135
QY 1599 GTGAGAGCGACACACCTGAGTGGCGGTGAAGTCTTGTGAGAGACGCTCCCACTGAC 1658
DB 2136 CTTGGGCTGATTAACAACAAGTGGCAGATTAAGAGCTGCGGAGACACATGCGCCGACAGC 2195
QY 1659 CTCAAGCCCAATTTTCTACAGAGCAGAGATCTGGAAGCAGATTAAGCCACCCCAATC 1718
DB 2196 TTAAAGGATTAATTTCTTAATGAGAGCAGAAATATGAAACAAATATAGTCACTTAACAT 2255
QY 1719 GTGCTGTCTATGTGTGACACCCAGAGAGCCATCTACATCTGTATGAGACTTGTG 1778

```

```
Db      2256 GTCAAACTTATTGAGTCTGCACAAAACATCCATTTAATTGTTATGAGACTAGTG 2315
Qy      1779 CAGGGGGGCGACTTCTCTGACCTTCTCTCGCAGAGAGGGGCCCGCTGCGGTGAAGACT 1838
Db      2316 CAAGGTGGAGACTTTTCAGACGTTTCTACAGAAATGAAGGTCCAGGCTGAAGGTGAAGAAG 2375
Qy      1839 CTGCTGCAGATGATGGGGGATGCAAGTGTCTGGCATGGAATACCTGGAGAGCAAGTCTGC 1898
Db      2376 CTTATTAGGGTGTCAAGAAATCTGCGGCTGGAATGGAATTTGGAGATTAACATTGT 2435
Qy      1899 ATCCACGGGACCTGAGCTGCTGCGAACTGCTGTGACAGAGAAAGATGCTCGAAGATC 1958
Db      2436 ATTCACAGGAGACTTGTGACAGCCCGTAACTGTCTGTGTCATGAGAAAATGCTTTAAAGATA 2495
Qy      1959 AGTGACTTTGGATGTCCCGAGAGAAAGCCGATGGGGTCTATGCAAGCTCAGGGGGCTTC 2018
Db      2496 AGTGACTTTGGATGTCCAGAGAAAGAAAGATGGAGTGTATTCATCCACTGGGGGGATG 2555
Qy      2019 AACAAGTCCCGTGAAGTGAACCGCACTGAAGCCCTTAACTAAGCCGCTACTCTCC 2078
Db      2556 AAGCAGATTCTTATTAAATGACGCGCCAGAAAGCCCTCAATTAACGTGCTACAGTTCA 2615
Qy      2079 GAAAGGACGTGAGAGCTTTGGCATCTTCTGCGGAGACCTTCAGCCTGGGGGCTCC 2138
Db      2616 GAGAGTGAATTTGGAGTTTGGATCTTAATCTGGAAGCCCTTCAGCTTGGGTTCTGTT 2675
Qy      2139 CCTATCCCACTCAGCAATCAGCAGACACGAGATTGTGAGAAAGGGGGCCGTCGTG 2198
Db      2676 CTTATGAGCTATGACAAACAGCAAAACAGGAGGCCATAGAACAGGAGTTGGTCTG 2735
Qy      2199 CCCTGCCAGAGCTGTCTCTGATGCTGTTCAAGGCTCATGAGCAGTCTGGGCTTAT 2258
Db      2736 CTTGTCCCGATPACTGCCCTGATGAAGTCTATAGCTGATGCTGAGATCTGGGAATAT 2795
Qy      2259 GAGCCTGGGCGAGGCCAGCTTCAGCACATCTAACAGAGCTGCAGAGCATCCGAAG 2318
Db      2796 GATCCAAAGAGGGGCTTAATTTCAGCATTTGCAATCAGGTGCTGTGCAACATCCGAAG 2855
Qy      2319 CGGCATCGGTGAGGCT 2334
Db      2856 AGATACCGATAAGTCT 2871
```

Search completed: March 19, 2005, 10:13:33
Job time : 1156.3 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 15:59:37 ; Search time 144 Seconds

(without alignments)
1726.159 Million cell updates/sec

Title: US-10-660-763-2

Perfect score: 3879

Sequence: 1 MGFSSELCSPOGHGVLQOMQ.....QRPSPSTIYQLSIRKXHR 752

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3879	100.0	752	US-10-003-295-2	Sequence 2, Appli
2	3879	100.0	752	US-10-660-763-2	Sequence 2, Appli
3	3834	98.8	822	US-10-003-295-4	Sequence 4, Appli
4	3834	98.8	822	US-10-660-763-4	Sequence 4, Appli
5	1853	47.8	822	US-09-886-319A-64	Sequence 64, Appli
6	1853	47.8	822	US-10-376-556-64	Sequence 64, Appli
7	1829.5	47.2	823	US-09-886-319A-63	Sequence 63, Appli
8	1829.5	47.2	823	US-10-376-556-63	Sequence 63, Appli
9	1528	39.4	292	US-10-664-421-69	Sequence 63, Appli
10	1444	37.2	275	US-10-763-418-38	Sequence 38, Appli
11	911	23.5	251	US-10-464-805-30	Sequence 30, Appli
12	899	23.2	251	US-08-987-689A-30	Sequence 30, Appli
13	899	23.2	251	US-10-292-524-30	Sequence 30, Appli

14	669.5	17.3	257	10	US-09-823-187-46	Sequence 46, Appli
15	668.5	17.2	257	10	US-09-863-776-42	Sequence 42, Appli
16	661	17.0	258	11	US-09-964-956-75	Sequence 75, Appli
17	661	17.0	258	15	US-10-042-865-100	Sequence 100, App
18	661	17.0	258	15	US-10-072-012-801	Sequence 801, App
19	661	17.0	258	15	US-10-072-012-858	Sequence 858, App
20	657	16.9	256	10	US-09-976-782-40	Sequence 40, Appli
21	657	16.9	257	10	US-09-976-782-29	Sequence 29, Appli
22	657	16.9	257	15	US-10-087-684-100	Sequence 100, App
23	657	16.9	257	15	US-10-218-779-100	Sequence 100, App
24	656	16.9	254	15	US-10-052-648A-43	Sequence 43, Appli
25	655	16.9	2029	16	US-10-746-545-33	Sequence 33, Appli
26	642	16.6	250	15	US-10-074-978A-153	Sequence 153, App
27	642	16.6	280	9	US-09-515-806-15	Sequence 15, Appli
28	642	16.6	280	15	US-10-423-543-17	Sequence 17, Appli
29	624.5	16.1	1130	14	US-10-204-041-4	Sequence 4, Appli
30	624.5	16.1	1148	15	US-10-170-385-181	Sequence 181, App
31	624.5	16.1	1148	16	US-10-408-765A-746	Sequence 746, App
32	623.5	16.1	278	14	US-10-172-088-9	Sequence 9, Appli
33	623	16.1	244	15	US-10-042-865-85	Sequence 85, Appli
34	623	16.1	1123	16	US-10-746-545-10	Sequence 10, Appli
35	623	16.1	1123	16	US-10-746-545-11	Sequence 11, Appli
36	623	16.1	1149	15	US-10-457-954-6	Sequence 6, Appli
37	622	16.0	1130	14	US-10-171-889-1	Sequence 1, Appli
38	622	16.0	1130	14	US-10-276-633-4	Sequence 4, Appli
39	622	16.0	1130	14	US-10-263-480-2	Sequence 2, Appli
40	620.5	16.0	1567	14	US-10-312-918-2	Sequence 2, Appli
41	620.5	16.0	1594	14	US-10-312-918-4	Sequence 4, Appli
42	618	15.9	343	16	US-10-664-421-57	Sequence 57, Appli
43	614.5	15.8	537	16	US-10-746-545-12	Sequence 12, Appli
44	614.5	15.8	537	16	US-10-746-545-13	Sequence 13, Appli
45	614.5	15.8	537	16	US-10-746-545-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-10-003-295-2
; Sequence 2, Application US/10003295
; Publication No. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; FILE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001183DIY
; CURRENT APPLICATION NUMBER: US/10/003, 295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 752
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-295-2

Query Match 100.0%; Score 3879; DB 13; Length 752;
Best Local Similarity 100.0%; Pred. No. 2.8e-223;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGFSSELCSPOGHGVLQOMQOEALRLLEGRKMAQORVKSDERYAGLHHMSLDSCGGS	60
DB	1	MGFSSELCSPOGHGVLQOMQOEALRLLEGRKMAQORVKSDERYAGLHHMSLDSCGGS	60
QY	61	RAISPSPISQSWARTTSQTEGLSRLLRQAHADLNSGPIKSLIRERQOLRKYSEOM	120
DB	61	RAISPSPISQSWARTTSQTEGLSRLLRQAHADLNSGPIKSLIRERQOLRKYSEOM	120
QY	121	QOLQOELTTHSODIEKLKQYRALARDAQAQKRYQASKQDKDKAKQKTVRSIMKLF	180
DB	121	QOLQOELTTHSODIEKLKQYRALARDAQAQKRYQASKQDKDKAKQKTVRSIMKLF	180

QY 181 AHHNRVYLVGRAAQLHHQHHLPLGLRSLODLHEEMACILKEILOEYLEISLVODE 240
DB 181 AHHNRVYLVGRAAQLHHQHHLPLGLRSLODLHEEMACILKEILOEYLEISLVODE 240
QY 241 VVAIHREMAAAARIQPEAEYOGFLROYGAPDVPCCVTFDSLSLEEGPPELGELOLNE 300
DB 241 VVAIHREMAAAARIQPEAEYOGFLROYGAPDVPCCVTFDSLSLEEGPPELGELOLNE 300
QY 301 LVESVQHTLSTVTDLAVATEMVFRRQEMVTOLOQELANEBENTHPRERVOLGKROYL 360
DB 301 LVESVQHTLSTVTDLAVATEMVFRRQEMVTOLOQELANEBENTHPRERVOLGKROYL 360
QY 361 QEALQGLQVALCSQAKLOAQOELLQTKLEHLGPGEPBPVLLLODDHSTSSSEOREGGR 420
DB 361 QEALQGLQVALCSQAKLOAQOELLQTKLEHLGPGEPBPVLLLODDHSTSSSEOREGGR 420
QY 421 TPTLEILKSHISGIRPKFSNLYRLEGGFPISIPLLIDHLSTQOPLTKSGVVLHRAVP 480
DB 421 TPTLEILKSHISGIRPKFSNLYRLEGGFPISIPLLIDHLSTQOPLTKSGVVLHRAVP 480
QY 481 KKKWVLNHDVLVGEQIGRGNFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR1 540
DB 481 KKKWVLNHDVLVGEQIGRGNFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR1 540
QY 541 LKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLFLTEGARLRVKTLLQWGDAAAG 600
DB 541 LKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLFLTEGARLRVKTLLQWGDAAAG 600
QY 601 MEYLESKCCIHDLAARNCLVTEKNVLIKISDFGMSREBDGVYAASGGLROYVPVKMTABE 660
DB 601 MEYLESKCCIHDLAARNCLVTEKNVLIKISDFGMSREBDGVYAASGGLROYVPVKMTABE 660
QY 661 ALNYGRYSESDVWSFGILLMETFSLGASPYNLSNOQTRFEVEKGRLLPCBELCPDAVF 720
DB 661 ALNYGRYSESDVWSFGILLMETFSLGASPYNLSNOQTRFEVEKGRLLPCBELCPDAVF 720
QY 721 RLMEQCMAYEPGQRPFSFTIYOELQSIKRRHR 752
DB 721 RLMEQCMAYEPGQRPFSFTIYOELQSIKRRHR 752

RESULT 2
US-10-660-763-2
; Sequence 2, Application US/10660763
; Publication No. US20040063130A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183DIYI
; CURRENT APPLICATION NUMBER: US/10/660, 763
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 752
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-660-763-2

Query Match 100.0%; Score 3879; DB 15; Length 752;
Best Local Similarity 100.0%; Pred. No. 2,8e-223;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSSSELCSPOGHGVLQOMQEAELRLLEGMRKMAORVXSDBREYAGLLHMSLODSGGOS 60
DB 1 MGSSSELCSPOGHGVLQOMQEAELRLLEGMRKMAORVXSDBREYAGLLHMSLODSGGOS 60
QY 61 RAISPSPIISQSAEITSGTEGSLRLRQHAEDLNSGPKSLLLRROQLKATYSQW 120
DB 61 RAISPSPIISQSAEITSGTEGSLRLRQHAEDLNSGPKSLLLRROQLKATYSQW 120

QY 121 QOLOELTKTSHQDIEKLSQYRALARDSAQAKRYQOASXKDRDKAKDYKVSIMKLF 180
DB 121 QOLOELTKTSHQDIEKLSQYRALARDSAQAKRYQOASXKDRDKAKDYKVSIMKLF 180
QY 181 AHHNRVYLVGRAAQLHHQHHLPLGLRSLODLHEEMACILKEILOEYLEISLVODE 240
DB 181 AHHNRVYLVGRAAQLHHQHHLPLGLRSLODLHEEMACILKEILOEYLEISLVODE 240
QY 241 VVAIHREMAAAARIQPEAEYOGFLROYGAPDVPCCVTFDSLSLEEGPPELGELOLNE 300
DB 241 VVAIHREMAAAARIQPEAEYOGFLROYGAPDVPCCVTFDSLSLEEGPPELGELOLNE 300
QY 301 LVESVQHTLSTVTDLAVATEMVFRRQEMVTOLOQELANEBENTHPRERVOLGKROYL 360
DB 301 LVESVQHTLSTVTDLAVATEMVFRRQEMVTOLOQELANEBENTHPRERVOLGKROYL 360
QY 361 QEALQGLQVALCSQAKLOAQOELLQTKLEHLGPGEPBPVLLLODDHSTSSSEOREGGR 420
DB 361 QEALQGLQVALCSQAKLOAQOELLQTKLEHLGPGEPBPVLLLODDHSTSSSEOREGGR 420
QY 421 TPTLEILKSHISGIRPKFSNLYRLEGGFPISIPLLIDHLSTQOPLTKSGVVLHRAVP 480
DB 421 TPTLEILKSHISGIRPKFSNLYRLEGGFPISIPLLIDHLSTQOPLTKSGVVLHRAVP 480
QY 481 KKKWVLNHDVLVGEQIGRGNFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR1 540
DB 481 KKKWVLNHDVLVGEQIGRGNFGEVFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR1 540
QY 541 LKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLFLTEGARLRVKTLLQWGDAAAG 600
DB 541 LKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLFLTEGARLRVKTLLQWGDAAAG 600
QY 601 MEYLESKCCIHDLAARNCLVTEKNVLIKISDFGMSREBDGVYAASGGLROYVPVKMTABE 660
DB 601 MEYLESKCCIHDLAARNCLVTEKNVLIKISDFGMSREBDGVYAASGGLROYVPVKMTABE 660
QY 661 ALNYGRYSESDVWSFGILLMETFSLGASPYNLSNOQTRFEVEKGRLLPCBELCPDAVF 720
DB 661 ALNYGRYSESDVWSFGILLMETFSLGASPYNLSNOQTRFEVEKGRLLPCBELCPDAVF 720
QY 721 RLMEQCMAYEPGQRPFSFTIYOELQSIKRRHR 752
DB 721 RLMEQCMAYEPGQRPFSFTIYOELQSIKRRHR 752

RESULT 3
US-10-003-295-4
; Sequence 4, Application US/10003295
; Publication No. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183DIYI
; CURRENT APPLICATION NUMBER: US/10/003, 295
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-295-4

Query Match 98.8%; Score 3834; DB 13; Length 822;
Best Local Similarity 91.5%; Pred. No. 1,5e-220;
Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGSSSELCSPOGHGVLQOMQEAELRLLEGMRKMAORVXSDBREYAGLLHMSLODSGGOS 60
DB 1 MGSSSELCSPOGHGVLQOMQEAELRLLEGMRKMAORVXSDBREYAGLLHMSLODSGGOS 60


```

QY 61 RAISPDSPISQSWAEITTSOTEGSLRLRQHAEDLNSGPKSLRLRERQOLKTYSEOW 120
DB 61 RAISPDSPISQSWAEITTSOTEGSLRLRQHAEDLNSGPKSLRLRERQOLKTYSEOW 120
QY 121 QOLQOELTTHSGDIEKLKSOYRALARDSQAKRKYQASXKODRKADKXVRSIMKLF 180
DB 121 QOLQOELTTHSGDIEKLKSOYRALARDSQAKRKYQASXKODRKADKXVRSIMKLF 180
QY 181 AHNHRYVLGVRAAQLHQNHHQHLPLGLRLSLQDLHEBNACILKEILOEYLEISLVODE 240
DB 181 AHNHRYVLGVRAAQLHQNHHQHLPLGLRLSLQDLHEBNACILKEILOEYLEISLVODE 240
QY 241 VVAIHREMAAARARIOPEAEYOGFLRQYGSADVPVPCVTFDSLEEGEPLEBGELOLNE 300
DB 241 VVAIHREMAAARARIOPEAEYOGFLRQYGSADVPVPCVTFDSLEEGEPLEBGELOLNE 300
QY 301 LTVESVQHTLTSTVDELAVATEMVFRRQEMVTQLOOELNBEENTHPRERVOQLGKROYL 360
DB 301 LTVESVQHTLTSTVDELAVATEMVFRRQEMVTQLOOELNBEENTHPRERVOQLGKROYL 360
QY 361 QEALQGLQVALCSQAQLQAOQELQTKLHLGPEBPVULLQDDHSTSSSEOREGGR 420
DB 361 QEALQGLQVALCSQAQLQAOQELQTKLHLGPEBPVULLQDDHSTSSSEOREGGR 420
QY 421 TPTLEILKSHISGIFRPKFS----- 440
DB 421 TPTLEILKSHISGIFRPKFS----- 440
QY 441 -----NLRYLEGEFPSPILPLIDHLLSTQOPLTKK 470
DB 441 LVBESGKQOEYVLSVMDGLPRHFIITQSLDNLKRLGEBGPSPILPLIDHLLSTQOPLTKK 540
QY 471 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFSGRLRADNTLVAAKSCRETLPPDL 530
DB 471 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFSGRLRADNTLVAAKSCRETLPPDL 530
QY 541 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFSGRLRADNTLVAAKSCRETLPPDL 600
DB 541 KAKFLQEARILKQYSHPNIVRLIGVCTQKOPYIYVWELVQGGDFLFLRTGEGARLKVTL 660
QY 591 LQWVGDAAGMEYLEBCKCIHRDLAARNCLVTEKNVLTISDFGMSREBAGVYAAAGGLR 720
DB 591 LQWVGDAAGMEYLEBCKCIHRDLAARNCLVTEKNVLTISDFGMSREBAGVYAAAGGLR 720
QY 651 QVPVKMTAPEALNYGRYSSESVDVWSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 710
DB 651 QVPVKMTAPEALNYGRYSSESVDVWSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 710
QY 721 QVPVKMTAPEALNYGRYSSESVDVWSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 780
DB 721 QVPVKMTAPEALNYGRYSSESVDVWSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 780
QY 711 CPELCPDAVPRIMEQCWAYEPGQRPSPSTTYOELQSIIRRRHR 822
DB 711 CPELCPDAVPRIMEQCWAYEPGQRPSPSTTYOELQSIIRRRHR 822

RESULT 4
US-10-660-763-4
; Sequence 4, Application US/10660763
; Publication No. US20040063130A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183DIV1
; CURRENT APPLICATION NUMBER: US/10/660, 763
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PaasSeq for Windows Version 4.0
; SEQ ID NO: 4
; LENGTH: 822
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-660-763-4

```

```

Query Match      98.8%; Score 3834; DB 15; Length 822;
Best Local Similarity 91.5%; Pred. No. 1.5e-220;
Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGFSSSELCPQHGVLQOQWQEAELRLLEGARKMAQRYVSDREYAGLIHHMSIQDSGGS 60
DB 1 MGFSSSELCPQHGVLQOQWQEAELRLLEGARKMAQRYVSDREYAGLIHHMSIQDSGGS 60
QY 61 RAISPDSPISQSWAEITTSOTEGSLRLRQHAEDLNSGPKSLRLRERQOLKTYSEOW 120
DB 61 RAISPDSPISQSWAEITTSOTEGSLRLRQHAEDLNSGPKSLRLRERQOLKTYSEOW 120
QY 121 QOLQOELTTHSGDIEKLKSOYRALARDSQAKRKYQASXKODRKADKXVRSIMKLF 180
DB 121 QOLQOELTTHSGDIEKLKSOYRALARDSQAKRKYQASXKODRKADKXVRSIMKLF 180
QY 181 AHNHRYVLGVRAAQLHQNHHQHLPLGLRLSLQDLHEBNACILKEILOEYLEISLVODE 240
DB 181 AHNHRYVLGVRAAQLHQNHHQHLPLGLRLSLQDLHEBNACILKEILOEYLEISLVODE 240
QY 241 VVAIHREMAAARARIOPEAEYOGFLRQYGSADVPVPCVTFDSLEEGEPLEBGELOLNE 300
DB 241 VVAIHREMAAARARIOPEAEYOGFLRQYGSADVPVPCVTFDSLEEGEPLEBGELOLNE 300
QY 301 LTVESVQHTLTSTVDELAVATEMVFRRQEMVTQLOOELNBEENTHPRERVOQLGKROYL 360
DB 301 LTVESVQHTLTSTVDELAVATEMVFRRQEMVTQLOOELNBEENTHPRERVOQLGKROYL 360
QY 361 QEALQGLQVALCSQAQLQAOQELQTKLHLGPEBPVULLQDDHSTSSSEOREGGR 420
DB 361 QEALQGLQVALCSQAQLQAOQELQTKLHLGPEBPVULLQDDHSTSSSEOREGGR 420
QY 421 TPTLEILKSHISGIFRPKFS----- 440
DB 421 TPTLEILKSHISGIFRPKFS----- 440
QY 441 -----NLRYLEGEFPSPILPLIDHLLSTQOPLTKK 470
DB 441 LVBESGKQOEYVLSVMDGLPRHFIITQSLDNLKRLGEBGPSPILPLIDHLLSTQOPLTKK 540
QY 471 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFSGRLRADNTLVAAKSCRETLPPDL 530
DB 471 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFSGRLRADNTLVAAKSCRETLPPDL 530
QY 541 SGVYLHRAVPKQKMWLNHEDVLGEOIGRNGEVSFSGRLRADNTLVAAKSCRETLPPDL 600
DB 541 KAKFLQEARILKQYSHPNIVRLIGVCTQKOPYIYVWELVQGGDFLFLRTGEGARLKVTL 660
QY 591 LQWVGDAAGMEYLEBCKCIHRDLAARNCLVTEKNVLTISDFGMSREBAGVYAAAGGLR 720
DB 591 LQWVGDAAGMEYLEBCKCIHRDLAARNCLVTEKNVLTISDFGMSREBAGVYAAAGGLR 720
QY 651 QVPVKMTAPEALNYGRYSSESVDVWSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 710
DB 651 QVPVKMTAPEALNYGRYSSESVDVWSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 710
QY 721 QVPVKMTAPEALNYGRYSSESVDVWSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 780
DB 721 QVPVKMTAPEALNYGRYSSESVDVWSFGILLMETFSLGASPYPLSNQOTREFVEKGRRLP 780
QY 711 CPELCPDAVPRIMEQCWAYEPGQRPSPSTTYOELQSIIRRRHR 822
DB 711 CPELCPDAVPRIMEQCWAYEPGQRPSPSTTYOELQSIIRRRHR 822

RESULT 5
US-09-886-319A-64
; Sequence 64, Application US/09886319A
; Publication No. US20020086019A1
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Weimer, Sabine
; APPLICANT: Halle, Jorn-Peter
; APPLICANT: Regembogen, Johannes
; APPLICANT: Goppelt, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound

```

```

/ TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
/ FILE OF INVENTION: Active Substances
/ FILE REFERENCE: 50125/014002
/ CURRENT APPLICATION NUMBER: US/09/886,319A
/ CURRENT FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: US 60/222,081
/ PRIOR FILING DATE: 2000-08-01
/ PRIOR APPLICATION NUMBER: DE 10030149.5
/ PRIOR FILING DATE: 2000-06-20
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: FaSTSeq for Windows Version 4.0
/ SEQ ID NO 64
/ LENGTH: 822
/ TYPE: PR
/ ORGANISM: Homo sapiens
US-09-886-319A-64

Query Match      47.8%; Score 1853; DB 9; Length 822;
Best Local Similarity 45.5%; Pred. No. 2,6e-102;
Matches 376; Conservative 150; Mismatches 217; Indels 84; Gaps 9;

QY 1 MGFSSSLCPBGHGVLCQOQOAEFLRLLEGMRKMAQRYVSDREYAGLLHMSLQDSGGOS 60
DB 1 MGFSSDL--KNSHEAVLKLQDWELRLLETVKFMALRISDKEXASTLQNLQNVDKEST 58
QY 61 RAISPDSPISQSWAETTSOTEGSLRLRQHADLNSGPISKLILRERQOLAKTYSEQM 120
DB 59 VQNALVSNVSKSWMLMIQTEQLSRIMKTHAEDLNSGPLRLTMMKDKQVKKSTYGVN 118
QY 121 QOLQOELTHTSODIEKLKSQYRALARDSAQAKRYQEA-SKODRKADKXYVSLMTL 179
DB 119 QOLIEAMIKVTTELEKLKCSYROLIKEMNSAKERYKALAKGTEKAKERYDKATMKT 178
QY 180 FAHNRVYLVGVAQAOHLQHHNOLLPLGLLSLODHEEMACILKEILOEYELISLVOD 239
DB 179 HMLHQYVLTALKAQDLHQYQYDITPLRLDLSQKQOEBIMALKGIFDEYSQITSLYTE 238
QY 240 EVVALIREMAAARLQPEAEYQGLRQYGSAPDVPCPTFDESLLEGPELPELQJN 299
DB 239 EIVNHNKEIQMSVEQIDPSTEYNNFIDVHRTTAAKQOELEFDTLSLEENENLOANEIMN 298
QY 300 ELTVESVQHTLTSVDELAVALATEMFRROEMVTQLOELRNEENTHPR-ERVOLLGKRQ 358
DB 299 NITASLQVWLKTLAELMQOQMLNKEAVLELEKRIEESSETECKSDIVLLSQKQ 358
QY 359 VLOEALQGLQVALCSQAKLQAOQELLQTKLEHLGPEPPVLLLODRHSTSSSEQERBG 418
DB 359 ALEELKQSVQQLRCTEAKFSQKELLEQKVQENDGKEPPVNVYEDASVTSMERKE-- 416
QY 419 GRTPTLEILKSHISGIFR-PK----- 438
DB 417 -RLSKFESTIRHSIAGIIRSPKSAVGSALSALSDMISSEKPLAEDQWYHGAIPIRLEAOBLK 475
QY 439 -----FSNLRYLLEGGEPPSIPLLDHLLSTQ 464
DB 476 KQGDPLVRSHGKRGGEVYLVSDGQRHPIIQVYNNMFRFEGTGSNIPQLDHHYTTK 535
QY 465 QPITKSGVYVLAHVARD-KVTLNHEDELVLGEQIGANGGEVPSGLRADNTLVAYKSCR 523
DB 536 QVTTKSGVYVLAHVARD-KVTLNHEDELVLGEQIGANGGEVPSGLRADNTLVAYKSCR 594
QY 524 ETLRPLKAKFLQOEAELIKQYSHNIVRLIGVCTOKOPITYIWEILOVGGPFLFTEG 583
DB 595 EDLPQELKFLQOEAELIKQYSHNIVRLIGVCTOKOPITYIWEILOVGGPFLFTEG 654
QY 584 RLKRVKTLQWVGDAAGMEYLESKCCIRHDLAARNCLVTEKNVLTSDFGMSHEADGVY 643
DB 655 ELKRLQVLRKSLDAAGMLYLESKCCIRHDLAARNCLVTEKNVLTSDFGMSHEADGVY 714
QY 644 AASGGLRQVYVKTAAABALNYGRYSSESDVWSFGIILMTFSGASPYRPLNSQOTREY 703
DB 715 SSS-GLKOIPIKWTABALNYGRYSSESDVWSFGIILMTFSGASPYRPLNSQOTREY 773

```

```

QY 704 EKGRLPCPELCPDAVFRLEMQWABPQORPSFTYIYQELQISRRK 750
DB 774 ERGYMSAPQHCPEDISIKIMKCMWYKPPENRRKFSLOKELITIKK 820

RESULT 6
US-10-376-564-64
/ Sequence 64, Application US/10376564
/ Publication No. US20030180302A1
/ GENERAL INFORMATION:
/ APPLICANT: Wolf, Eckhard
/ APPLICANT: Werner, Sabine
/ APPLICANT: Haile, Jörn-Peter
/ APPLICANT: Regenhagen, Johannes
/ APPLICANT: Goppelt, Andreas
/ TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
/ TITLE OF INVENTION: the diagnosis or treatment of skin disorders and wound
/ TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
/ FILE REFERENCE: 50125/014003
/ CURRENT APPLICATION NUMBER: US/10/376,564
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 09/886,319
/ PRIOR FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: US 60/222,081
/ PRIOR FILING DATE: 2000-08-01
/ PRIOR APPLICATION NUMBER: DE 100 30 149.5-41
/ PRIOR FILING DATE: 2000-06-20
/ NUMBER OF SEQ ID NOS: 86
/ SOFTWARE: FaSTSeq for Windows Version 4.0
/ SEQ ID NO 64
/ LENGTH: 822
/ TYPE: PR
/ ORGANISM: Homo sapiens
US-10-376-564-64

Query Match      47.8%; Score 1853; DB 14; Length 822;
Best Local Similarity 45.5%; Pred. No. 2,6e-102;
Matches 376; Conservative 150; Mismatches 217; Indels 84; Gaps 9;

QY 1 MGFSSSLCPBGHGVLCQOQOAEFLRLLEGMRKMAQRYVSDREYAGLLHMSLQDSGGOS 60
DB 1 MGFSSDL--KNSHEAVLKLQDWELRLLETVKFMALRISDKEXASTLQNLQNVDKEST 58
QY 61 RAISPDSPISQSWAETTSOTEGSLRLRQHADLNSGPISKLILRERQOLAKTYSEQM 120
DB 59 VQNALVSNVSKSWMLMIQTEQLSRIMKTHAEDLNSGPLRLTMMKDKQVKKSTYGVN 118
QY 121 QOLQOELTHTSODIEKLKSQYRALARDSAQAKRYQEA-SKODRKADKXYVSLMTL 179
DB 119 QOLIEAMIKVTTELEKLKCSYROLIKEMNSAKERYKALAKGTEKAKERYDKATMKT 178
QY 180 FAHNRVYLVGVAQAOHLQHHNOLLPLGLLSLODHEEMACILKEILOEYELISLVOD 239
DB 179 HMLHQYVLTALKAQDLHQYQYDITPLRLDLSQKQOEBIMALKGIFDEYSQITSLYTE 238
QY 240 EVVALIREMAAARLQPEAEYQGLRQYGSAPDVPCPTFDESLLEGPELPELQJN 299
DB 239 EIVNHNKEIQMSVEQIDPSTEYNNFIDVHRTTAAKQOELEFDTLSLEENENLOANEIMN 298
QY 300 ELTVESVQHTLTSVDELAVALATEMFRROEMVTQLOELRNEENTHPR-ERVOLLGKRQ 358
DB 299 NITASLQVWLKTLAELMQOQMLNKEAVLELEKRIEESSETECKSDIVLLSQKQ 358
QY 359 VLOEALQGLQVALCSQAKLQAOQELLQTKLEHLGPEPPVLLLODRHSTSSSEQERBG 418
DB 359 ALEELKQSVQQLRCTEAKFSQKELLEQKVQENDGKEPPVNVYEDASVTSMERKE-- 416
QY 419 GRTPTLEILKSHISGIFR-PK----- 438
DB 417 -RLSKFESTIRHSIAGIIRSPKSAVGSALSALSDMISSEKPLAEDQWYHGAIPIRLEAOBLK 475
QY 439 -----FSNLRYLLEGGEPPSIPLLDHLLSTQ 464

```

```

Db 476 KQGDPLVRESHGKPGBYLVSVDGQRHFIIOYDMMTRFEGTGSNIPOLIDHHYTK 535
Qy 465 QPLTKKSGVLIHRAVPKD-KVNLNHEDVLGEQIGRNGFGEVFSGLRADNTLVAVKSC 523
Db 536 QVTTKSGVLIHRAVPKD-KVNLNHEDVLGEQIGRNGFGEVFSGLRADNTLVAVKSC 594
Qy 524 ETLPPDLKAKFLOEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFLRTGA 583
Db 595 EDLPQELKIKFLOEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFLRTGA 654
Qy 584 RLKRVKLIOMVGDAAGMEYLESKCIHRDLAARNCILTEKNVLKISDPGMSREBAGYV 643
Db 655 ELKQVLKQVFSIDPAAGMILYESKNCIHRDLAARNCILTEKNVLKISDPGMSREBAGYV 714
Qy 644 AASGGLQVVPVMTABEALNRYGSSSDVWSPGILLMETFSLGASPPYPLNSNOQTRFV 703
Db 715 SSS-GLKQIPIKWTADEALNRYGSSSDVWSPGILLMETFSLGASPPYPLNSNOQTRFV 773
Qy 704 EKGGRLPCEBELCPDAVFRLMEOCMAYEPGQRPSEFTIYQELQSIK 750
Db 774 EKGGRMSAPQNCPEEVFTIMKCMQDYKPNRPFSLOKELTIK 820

```

RESULT 7

```

US-09-886-319A-63
/ Sequence 63, Application US/09886319A
/ Publication No. US20020086019A1
/ GENERAL INFORMATION:
/ APPLICANT: Wolf, Eckard
/ APPLICANT: Werner, Sabine
/ APPLICANT: Halle, Jörn-Peter
/ APPLICANT: Regenbogen, Johannes
/ APPLICANT: Goppelt, Andreas
/ TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
/ TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
/ TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
/ TITLE OF INVENTION: Active Substances
/ FILE REFERENCE: 50125/014002
/ CURRENT APPLICATION NUMBER: US/09/886,319A
/ PRIOR FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: US 60/222,081
/ PRIOR FILING DATE: 2000-08-01
/ PRIOR APPLICATION NUMBER: DE 10030149.5
/ PRIOR FILING DATE: 2000-06-20
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 63
/ LENGTH: 823
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-09-886-319A-63

```

```

Query Match 47.2%, Score 1829.5; DB 9; Length 823;
Best Local Similarity 45.6%, Pred. No. 6.5e-101;
Matches 377; Conservative 146; Mismatches 219; Indels 85; Gaps 10;

```

```

Qy 1 MGFSSELCSPOGHVLIQWQOAEILRLLEGKMKMAORVSDREYAGILHMSLQDSGGS 60
Db 1 MEGSGSLKNSQ-BAYL-KIQDWELRLLEFYKKFMARISDKEXAYATLQNLQNVKEST 58
Qy 61 RAISPSPIQSQAETSOTEGSLRLRQDAEDLNSGRLSKSLIRERQQLKTYSEGM 120
Db 59 VQVNVYVNSKSVLMIQOTEOELSRIMKTAEDLNSGRLHRLTMKIKDQVKKSYVGIH 118
Qy 121 OOLQOELTTHSDIKLSQVYALARDSAQAKRYQEA-SXQKDXDKAKDXVRSMLK 179
Db 119 QOIBAMIKVTKTELEKLSYRQLTKENMSAKETKEALAKETEKAKERDKATMK 178
Qy 180 FAHNNYLVGVRAAQHNOHQLLPGILRSIQDHEEMACILKEILOEYLISSLYVD 239
Db 179 HMLHNGVYALAKGAQHOGQYVDTTLPLLLDSVQKQOEBMIALKQIPDDYSQITSLVTE 238

```

```

Qy 240 EVVAIHRMAAARIAIOPEAEYQGFLOXGASBPVPCVTFPESELSEBEPLEPGLQLN 299
Db 239 EIVNHKEIQNSVEQIDPSTETNNFIDVARTTAKEQIEIFDLSLEENENIQANEMNN 298
Qy 300 ELTVESQHTLTSVTDLAVATEMFRROEMVTOLOEELRNEENHPR-ERVOLLGKRO 358
Db 299 NLTAADSLQVMTLTIABELQTOQMLHKAVALLEKRIEESPECEKSDVILLGQNG 358
Qy 359 VLQELQIGVALCSQAKIQAOQELLQTKLBNLGRPEPPVILLDDNRHSTSSSEQERBG 418
Db 359 ALBELKQSVQQRCSAKCAQKALLEQVQENDGKEPPVNVYNEEDARSVTSMEKKE-- 416
Qy 419 GRTPLTEIKSHISGFR-PK----- 438
Db 417 -RLSKFESTIRHSIAGITKSPKSVLQSSITQCVISVGERPLAEHDMTHCAIRIEAQELL 475
Qy 439 -----FSNLYRLEGEGFPSPILILDLHLSLT 463
Db 476 KQGDPLVRESHGKPGBYLVSVDGQRHFIIOYDMMTRFEGTGSNIPOLIDHHYTK 535
Qy 464 QPLTKKSGVLIHRAVPKD-KVNLNHEDVLGEQIGRNGFGEVFSGLRADNTLVAVKSC 522
Db 536 QVTTKSGVLIHRAVPKD-KVNLNHEDVLGEQIGRNGFGEVFSGLRADNTLVAVKSC 594
Qy 523 RETLPPDLKAKFLOEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFLRTGA 582
Db 595 EDLPQELKIKFLOEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFLRTGA 654
Qy 584 RLKRVKLIOMVGDAAGMEYLESKCIHRDLAARNCILTEKNVLKISDPGMSREBAGYV 642
Db 655 ELKQVLKQVFSIDPAAGMILYESKNCIHRDLAARNCILTEKNVLKISDPGMSREBAGYV 714
Qy 643 YVAGSGLQVVPVMTABEALNRYGSSSDVWSPGILLMETFSLGASPPYPLNSNOQTRFV 702
Db 715 YSSS-GLKQIPIKWTADEALNRYGSSSDVWSPGILLMETFSLGASPPYPLNSNOQTRFV 773
Qy 703 VKGGRLPCEBELCPDAVFRLMEOCMAYEPGQRPSEFTIYQELQSIK 749
Db 774 VERGRMSAPQNCPEEVFTIMKCMQDYKPNRPFSLOKELTIK 820

```

RESULT 8

```

US-10-376-564-63
/ Sequence 63, Application US/10376564
/ Publication No. US20030180302A1
/ GENERAL INFORMATION:
/ APPLICANT: Wolf, Eckard
/ APPLICANT: Werner, Sabine
/ APPLICANT: Halle, Jörn-Peter
/ APPLICANT: Regenbogen, Johannes
/ APPLICANT: Goppelt, Andreas
/ TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
/ TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
/ TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
/ TITLE OF INVENTION: Active Substances
/ FILE REFERENCE: 50125/014003
/ CURRENT APPLICATION NUMBER: US/10/376,564
/ PRIOR FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 09/886,319
/ PRIOR FILING DATE: 2001-06-20
/ PRIOR APPLICATION NUMBER: US 60/222,081
/ PRIOR FILING DATE: 2000-08-01
/ PRIOR APPLICATION NUMBER: DE 100 30 149.5-41
/ PRIOR FILING DATE: 2000-06-20
/ NUMBER OF SEQ ID NOS: 86
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 63
/ LENGTH: 823
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-10-376-564-63

```

```

Query Match 47.2%, Score 1829.5; DB 14; Length 823;

```

Best Local Similarity 45.6%; Pred. No. 6.5e-101;
Matches 377; Conservative 146; Mismatches 219; Indels 85; Gaps 10;

```

Qy 1 MGFSSSLCPQGHVUQOQZAEALLEGKMKMAQRYVSDERYAGLLHMSLQDSGGG 60
Db 1 MGFSGDLKNSQ-BAVL-KLQDWELRLLETVKKFMALRISDKRYAYATLONLCNQVDEST 58
Qy 61 RAISPDSPISQSWAEITSGEGSLRLRQHAEDLNSGPKSLILRERQOLRKYSEGM 120
Db 59 VQNVVSNVSNVSKSLMLIQOTEOLSRIMKTHAEDLNSGPKSLILRERQOLRKYSEGM 118
Qy 121 QOLQOELTHTSHQDIEKLKSOYRALARDQAQKRYQEA-SKDQDKAKDKYVSLMTL 179
Db 119 QQIEAMKIVTTELEKLSRYQLIKEMNSAKETKALAKGETEKAKERYDKATMTL 178
Qy 180 FAHNRVYLGVAAQULHQQHQLLPGLIRSLQDHEMACLKEILOEYELISLVOD 239
Db 179 HMLHNOYVALKMAQOHQSOYDITPLRLIDSVOKQOEMIKALKGIFFDYSQITSLVTE 238
Qy 240 EYVALIHRMAAARIQPEAYOGFLROYGSAADVPDPCVFPDESLEEGEPLEPGELQIN 299
Db 239 EIVNVHKEIQMSVEQIDPSTETYNFIDVHRTTAAKQOEIEFDTSLSEENLQANETMNN 298
Qy 300 ELTVESVQHTLTSVDELAVATEMVERROBMVYTOLOQEDRNEBENTHPR-ERYOLLGKQ 358
Db 299 NLTADSLQVMTLKTALBELTQOQMLHKEAVLELEKRIEESFETCEKSDIVLLGKQ 358
Qy 359 VLOEALQGLQVALCSQAKIQAOQELIQTLEHGPBEPPVLLLODRSTSSSEBERG 418
Db 359 ALEELKQSVQOQRCSAKCAQVALLERQVQENDKEPPVNVYEDARSVYSMERKE-- 416
Qy 419 GRTPLTEIKSHISGIFR-PK----- 438
Db 417 -RLSKESIRHSIAGIISPKSVLSSYOVGVISGERPLAHDVYHGAIRIEKQELL 475
Qy 439 -----FSNLYRLEGESFPSILPLIDHLLST 463
Db 476 KQGGDFLVESHQKPEGVYLVSVSDGRRHFIIOFVDNLVRFEGTFSNIPQIDHFMNT 535
Qy 464 QOPLTKKSGVVLHRAVPKD-KVTLNHEDLVGEQIRGNFGEVFSRLADNTLVAVKSC 522
Db 536 KQVITTKSGVVLNLPKPKKWLNEHDVSLGELLKGNFGEVYKGTILK-DKTPVAILKTC 594
Qy 523 RETLPPDLKAKFLQEARILKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLTFRTG 582
Db 595 KEDLPBELIKFLQEARILKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLTFRTG 654
Qy 583 ARLRVKTLQMGVDAAGMEYLESKCCIHEDLAARNCLVTEKNVLSIDFGMSREBAGV 642
Db 655 DELKTLQVRFSLDVAAGMLYLESKCCIHEDLAARNCLVTEKNVLSIDFGMSREBAGV 714
Qy 643 YAAAGSLRQVPVWVTPALNLYGRYSESDVMSFGILLMETFSLGASPYPNLSNOQTRF 702
Db 715 YSSS-GLKQIPIMTLPALNLYGRYSESDVMSFGILLMETFSLGASPYPNLSNOQTRF 773
Qy 703 VEKGRGLPCPELCPDAVFRLEMCMAYPEQORSFSTIYQELQIRK 749
Db 774 VERGYMSAPQNCPEVFTIMKCMIDYKEBNRPKFDLKELTIVIK 820

```

RESULT 9
US-10-664-421-69

```

; Sequence 69, Application US/10664421
; Publication No. US2004012864A1
; GENERAL INFORMATION:
; APPLICANT: BREMER, RYAN
; APPLICANT: IBRAHIM, PRABHA
; APPLICANT: KUMAR, ABHINAV
; APPLICANT: MANDIYAN, VALSAN
; APPLICANT: MILBURN, MICHAEL V.
; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE
; FILE REFERENCE: 039363/0703
; CURRENT APPLICATION NUMBER: US/10/664,421

```

```

; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/412,341
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/411,398
; PRIOR FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 69
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-664-421-69

```

Query Match 39.4%; Score 1528; DB 16; Length 292;
Best Local Similarity 99.7%; Pred. No. 2e-83;
Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 461 LSTQOPLTKKSGVVLHRAVPKDQVTLNHEDLVLSQIGRGNFGEVFSGRRLRADNTLVAVK 520
Db 1 LSTQOPLTKKSGVVLHRAVPKDQVTLNHEDLVLSQIGRGNFGEVFSGRRLRADNTLVAVK 60
Qy 521 SRETLPPDLKAKFLQEARILKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLTFRT 580
Db 61 SRETLPPDLKAKFLQEARILKQYSHPNIVRLIGVCTOKOPYIYIMELVQGGDFLTFRT 120
Qy 581 EGARLRVKTLLQMGVDAAGMEYLESKCCIHEDLAARNCLVTEKNVLSIDFGMSREBAG 640
Db 121 EGARLRVKTLLQMGVDAAGMEYLESKCCIHEDLAARNCLVTEKNVLSIDFGMSREBAG 180
Qy 641 GYAAAGSLRQVPVWVTPALNLYGRYSESDVMSFGILLMETFSLGASPYPNLSNOQTR 700
Db 181 GYAAAGSLRQVPVWVTPALNLYGRYSESDVMSFGILLMETFSLGASPYPNLSNOQTR 240
Qy 701 EFKVGRGLPCPELCPDAVFRLEMCMAYPEQORSFSTIYQELQIRK 752
Db 241 EFKVGRGLPCPELCPDAVFRLEMCMAYPEQORSFSTIYQELQIRK 292

```

RESULT 10

```

US-10-763-418-38
; Sequence 38, Application US/10763418
; Publication No. US20040185547A1
; GENERAL INFORMATION:
; APPLICANT: MOHAMMADI, MOOSA
; APPLICANT: SCHLESSINGER, JOSEPH
; APPLICANT: HUBBARD, STEVAN R.
; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
; FILE REFERENCE: 038602/0847
; CURRENT APPLICATION NUMBER: US/10/763,418
; CURRENT FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: US/09/664,526
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 09/188,809
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 08/701,191
; PRIOR FILING DATE: 1996-08-21
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-763-418-38

```

Query Match 37.2%; Score 1444; DB 16; Length 275;
Best Local Similarity 99.6%; Pred. No. 1.9e-78;
Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 478 AVPKDKVTLNHEDLVLSQIGRGNFGEVFSGRRLRADNTLVAVKSGRETLPPDLKAKFLQ 537
Db 1 AVPKDKVTLNHEDLVLSQIGRGNFGEVFSGRRLRADNTLVAVKSGRETLPPDLKAKFLQ 60

```


This Page Blank (uspto)

Db 361 OEALQGLQVALCSQAQLQAQOELLQTKLEHLGPEBPPLLLQDDRHSTSSSEOREGGR 420
Qy 421 TPTLEILKSHISGIRFPKPSNLYRLEGEGFPSPILLIDHLSTQOPLTKKSGVYLHRAVP 480
Db 421 TPTLEILKSHISGIRFPKPSNLYRLEGEGFPSPILLIDHLSTQOPLTKKSGVYLHRAVP 480
Qy 481 KDKWVNLHEDLVLEGOIGRGNFGEVSGRLRADNTLVAAVSCRETIPLPKAKFLOEARI 540
Db 481 KDKWVNLHEDLVLEGOIGRGNFGEVSGRLRADNTLVAAVSCRETIPLPKAKFLOEARI 540
Qy 541 LKQYSHPNIVRLIGVCTOKOPIYIMELVQGGDFLFTLFTREGARLVKTLQWVGDAAG 600
Db 541 LKQYSHPNIVRLIGVCTOKOPIYIMELVQGGDFLFTLFTREGARLVKTLQWVGDAAG 600
Qy 601 MEYLESKCCIHRLAARNCLVTEKNVLKISDFGMSREBAGVYAASGGLRQVPVKWTABE 660
Db 601 MEYLESKCCIHRLAARNCLVTEKNVLKISDFGMSREBAGVYAASGGLRQVPVKWTABE 660
Qy 661 ALNYGRYSSESDVMSFGILLMETFSLGASPYPNLSNQOTREVEKGRPLPCBELCPDAVF 720
Db 661 ALNYGRYSSESDVMSFGILLMETFSLGASPYPNLSNQOTREVEKGRPLPCBELCPDAVF 720
Qy 721 RLMEQCMAYEPGQRPSPSTIYQELQSIKRRHR 752
Db 721 RLMEQCMAYEPGQRPSPSTIYQELQSIKRRHR 752

RESULT 2
US-10-003-295-2
; Sequence 2, Application US/10003295
; Patent No. 6686187
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 752
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-003-295-2

Query Match 100.0%; Score 3879; DB 4; Length 752;
Best Local Similarity 100.0%; Pred. No. 1.1e-261;
Matches 752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 301 LTVESVOHTLLSVTDELAVATEMVFROEMVTOLOQELRNEBENTHPRERVOILLGRQVYL 360
Qy 361 OEALQGLQVALCSQAQLQAQOELLQTKLEHLGPEBPPLLLQDDRHSTSSSEOREGGR 420
Db 361 OEALQGLQVALCSQAQLQAQOELLQTKLEHLGPEBPPLLLQDDRHSTSSSEOREGGR 420
Qy 421 TPTLEILKSHISGIRFPKPSNLYRLEGEGFPSPILLIDHLSTQOPLTKKSGVYLHRAVP 480
Db 421 TPTLEILKSHISGIRFPKPSNLYRLEGEGFPSPILLIDHLSTQOPLTKKSGVYLHRAVP 480
Qy 481 KDKWVNLHEDLVLEGOIGRGNFGEVSGRLRADNTLVAAVSCRETIPLPKAKFLOEARI 540
Db 481 KDKWVNLHEDLVLEGOIGRGNFGEVSGRLRADNTLVAAVSCRETIPLPKAKFLOEARI 540
Qy 541 LKQYSHPNIVRLIGVCTOKOPIYIMELVQGGDFLFTLFTREGARLVKTLQWVGDAAG 600
Db 541 LKQYSHPNIVRLIGVCTOKOPIYIMELVQGGDFLFTLFTREGARLVKTLQWVGDAAG 600
Qy 601 MEYLESKCCIHRLAARNCLVTEKNVLKISDFGMSREBAGVYAASGGLRQVPVKWTABE 660
Db 601 MEYLESKCCIHRLAARNCLVTEKNVLKISDFGMSREBAGVYAASGGLRQVPVKWTABE 660
Qy 661 ALNYGRYSSESDVMSFGILLMETFSLGASPYPNLSNQOTREVEKGRPLPCBELCPDAVF 720
Db 661 ALNYGRYSSESDVMSFGILLMETFSLGASPYPNLSNQOTREVEKGRPLPCBELCPDAVF 720
Qy 721 RLMEQCMAYEPGQRPSPSTIYQELQSIKRRHR 752
Db 721 RLMEQCMAYEPGQRPSPSTIYQELQSIKRRHR 752

RESULT 3
US-09-817-180-4
; Sequence 4, Application US/09817180
; Patent No. 6340584
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001183
; CURRENT APPLICATION NUMBER: US/09/817,180
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FaastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 822
; TYPE: PR1
; ORGANISM: Human
US-09-817-180-4

Query Match 98.8%; Score 3814; DB 3; Length 822;
Best Local Similarity 91.5%; Pred. No. 1.6e-258;
Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

Db 241 VVAIHREMAAARARLOPEAEYQGLRQYGSADVPVPCVTFDESLLEBGEPLRPGELQJNE 300
Qy 301 LTVESVQHTLTSVTDELAVATEMVFRRQEMVTLQOQELRNEBENTHPRRVLGKQRYL 360
Db 301 LTVESVQHTLTSVTDELAVATEMVFRRQEMVTLQOQELRNEBENTHPRRVLGKQRYL 360
Qy 361 QBALGGLQVALCSQAKLQAOQELQTKLEHLGPGEPPLVLLQODRHSTSSSEOREGGR 420
Db 361 QBALGGLQVALCSQAKLQAOQELQTKLEHLGPGEPPLVLLQODRHSTSSSEOREGGR 420
Qy 421 TPTLEILKSHISGIFRPFKS-----NLYRLEGEGRPSIPLLIDHLSTQOPLTKK 440
Db 421 TPTLEILKSHISGIFRPFKS-----NLYRLEGEGRPSIPLLIDHLSTQOPLTKK 440
Qy 481 LVRESQKQEVYLVLMQGLPRHFIIQSUDNLYRLEGEGRPSIPLLIDHLSTQOPLTKK 540
Db 471 SGVVLHRAVPKQKWLNHEDLVLEQIGRNGFGEVPSGRLRADNTLVAVKSCRETLRPLD 530
Qy 541 SGVVLHRAVPKQKWLNHEDLVLEQIGRNGFGEVPSGRLRADNTLVAVKSCRETLRPLD 600
Db 531 KAKPLQEARILKQYSHPNIVRLIGVCTOKOPITYIWMELVQGGDFLFLTEGARLRVKTL 590
Qy 601 KAKPLQEARILKQYSHPNIVRLIGVCTOKOPITYIWMELVQGGDFLFLTEGARLRVKTL 660
Db 591 LQWGDAAAGMEYLSKSCCIHRDLAARNCLVTEKNVLTISDFGMSREADGVYAAASGGLR 650
Qy 661 LQWGDAAAGMEYLSKSCCIHRDLAARNCLVTEKNVLTISDFGMSREADGVYAAASGGLR 720
Db 651 QVPVKMTAPBALNYGRYSSSDVMSFGILLMETFSLGASPYNLSNQOTREFVEKGRPL 710
Qy 721 QVPVKMTAPBALNYGRYSSSDVMSFGILLMETFSLGASPYNLSNQOTREFVEKGRPL 780
Db 711 CPELCPDVAFRLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 752
Qy 781 CPELCPDVAFRLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 822

RESULT 4
US-10-003-295-4
; Sequence 4, Application US/10003295
; Patent No. 6686187
; GENERAL INFORMATION:
; APPLICANT: GAN, weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-003-295-4

Query Match 98.8%; Score 3834; DB 4; Length 822;
Best Local Similarity 91.5%; Pred. No. 1.6e-258;
Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

Qy 1 MGFSSSELSPQGHGVLQMOQEAELRLLEGKRMMAQRYVSDREYAGLLHMSLQDSGQOS 60
Db 1 MGFSSSELSPQGHGVLQMOQEAELRLLEGKRMMAQRYVSDREYAGLLHMSLQDSGQOS 60
Qy 61 RAISPSPISQSWAETTSQTEGSLRLRQHAEDLNSGPIKSLILREHQQLKTYSEOW 120
Db 61 RAISPSPISQSWAETTSQTEGSLRLRQHAEDLNSGPIKSLILREHQQLKTYSEOW 120
Qy 121 QQLQOGLTHTHSQDIETKLSQYRALARSDAQAKRYQASXKQDRKAKDKYVRSIMKLF 180

Db 121 QQLQOGLTHTHSQDIETKLSQYRALARSDAQAKRYQASXKQDRKAKDKYVRSIMKLF 180
Qy 181 AHHNRYVAGVRAAQIHHQHHLPLGILRSIQDLHEBACTIKELQRYLEISLVODE 240
Db 181 AHHNRYVAGVRAAQIHHQHHLPLGILRSIQDLHEBACTIKELQRYLEISLVODE 240
Qy 241 VVAIHREMAAARARLOPEAEYQGLRQYGSADVPVPCVTFDESLLEBGEPLRPGELQJNE 300
Db 241 VVAIHREMAAARARLOPEAEYQGLRQYGSADVPVPCVTFDESLLEBGEPLRPGELQJNE 300
Qy 301 LTVESVQHTLTSVTDELAVATEMVFRRQEMVTLQOQELRNEBENTHPRRVLGKQRYL 360
Db 301 LTVESVQHTLTSVTDELAVATEMVFRRQEMVTLQOQELRNEBENTHPRRVLGKQRYL 360
Qy 361 QBALGGLQVALCSQAKLQAOQELQTKLEHLGPGEPPLVLLQODRHSTSSSEOREGGR 420
Db 361 QBALGGLQVALCSQAKLQAOQELQTKLEHLGPGEPPLVLLQODRHSTSSSEOREGGR 420
Qy 421 TPTLEILKSHISGIFRPFKS-----NLYRLEGEGRPSIPLLIDHLSTQOPLTKK 440
Db 421 TPTLEILKSHISGIFRPFKS-----NLYRLEGEGRPSIPLLIDHLSTQOPLTKK 440
Qy 481 LVRESQKQEVYLVLMQGLPRHFIIQSUDNLYRLEGEGRPSIPLLIDHLSTQOPLTKK 540
Db 471 SGVVLHRAVPKQKWLNHEDLVLEQIGRNGFGEVPSGRLRADNTLVAVKSCRETLRPLD 530
Qy 541 SGVVLHRAVPKQKWLNHEDLVLEQIGRNGFGEVPSGRLRADNTLVAVKSCRETLRPLD 600
Db 531 KAKPLQEARILKQYSHPNIVRLIGVCTOKOPITYIWMELVQGGDFLFLTEGARLRVKTL 590
Qy 601 KAKPLQEARILKQYSHPNIVRLIGVCTOKOPITYIWMELVQGGDFLFLTEGARLRVKTL 660
Db 591 LQWGDAAAGMEYLSKSCCIHRDLAARNCLVTEKNVLTISDFGMSREADGVYAAASGGLR 650
Qy 661 LQWGDAAAGMEYLSKSCCIHRDLAARNCLVTEKNVLTISDFGMSREADGVYAAASGGLR 720
Db 651 QVPVKMTAPBALNYGRYSSSDVMSFGILLMETFSLGASPYNLSNQOTREFVEKGRPL 710
Qy 721 QVPVKMTAPBALNYGRYSSSDVMSFGILLMETFSLGASPYNLSNQOTREFVEKGRPL 780
Db 711 CPELCPDVAFRLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 752
Qy 781 CPELCPDVAFRLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 822

RESULT 5
US-09-538-092-866
; Sequence 866, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Iolc
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 866
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P07332
US-09-538-092-866

```

Query Match      98.7%; Score 3828; DB 4; Length 822;
Best Local Similarity 91.4%; Pred. No. 4, 2e-258;
Matches 751; Conservative 0; Mismatches 1; Indels 70; Gaps 1;

QY 1 MGFSSSELCPGQGHVYLQOMQOAEELRLLEGMKMKMAQRYKSDREYAGLLHMSLQDSGGGS 60
DB 1 MGFSSSELCPGQGHVYLQOMQOAEELRLLEGMKMKMAQRYKSDREYAGLLHMSLQDSGGGS 60
QY 61 RAISPSPISSQSWAETTSQTEGSLRLRQHAEDLNSGPKSLSLIREQOLKRTYSEOW 120
DB 61 RAISPSPISSQSWAETTSQTEGSLRLRQHAEDLNSGPKSLSLIREQOLKRTYSEOW 120
QY 121 OQLOQELTKTHSODIKELKSYRALARDSAQAQKRYQEA KODKRDKADKRYSLMKLF 180
DB 121 OQLOQELTKTHSODIKELKSYRALARDSAQAQKRYQEA KODKRDKADKRYSLMKLF 180
QY 181 AHNHRYVLGVRAAQLHQHQQHQLLPGLRLSLQDLHEEMACILKEILQEYLEISSLVQDE 240
DB 181 AHNHRYVLGVRAAQLHQHQQHQLLPGLRLSLQDLHEEMACILKEILQEYLEISSLVQDE 240
QY 241 VVAIHREMAAARLQPEAEYQGFILQYGSAPVPCVTFDESLLBEGEPLEBGELOQNE 300
DB 241 VVAIHREMAAARLQPEAEYQGFILQYGSAPVPCVTFDESLLBEGEPLEBGELOQNE 300
QY 301 LTVESVQHTLTSYTDLAVALATEMVFROEMVTQLOQELRNEBENTHPRERVOILLGRQYL 360
DB 301 LTVESVQHTLTSYTDLAVALATEMVFROEMVTQLOQELRNEBENTHPRERVOILLGRQYL 360
QY 361 QEALQGLQVALCSQAQLQAQOELLQTKLEHLGEBEPVLLLODDRHSSTSSBOEREGGR 420
DB 361 QEALQGLQVALCSQAQLQAQOELLQTKLEHLGEBEPVLLLODDRHSSTSSBOEREGGR 420
QY 421 TPTLEILKSHISGIFRPKFS-----NLRYLGEGBPPSIPLIDHLLSTQOPLTKK 440
DB 421 TPTLEILKSHISGIFRPKFS-----NLRYLGEGBPPSIPLIDHLLSTQOPLTKK 440
QY 441 -----NLRYLGEGBPPSIPLIDHLLSTQOPLTKK 470
DB 441 -----NLRYLGEGBPPSIPLIDHLLSTQOPLTKK 470
QY 471 SGVVLHRAVPKDKMVLNHEBDLVLEQDIGRNGFGEVSGRLRADNTLVAVKSCRETLPPDL 530
DB 471 SGVVLHRAVPKDKMVLNHEBDLVLEQDIGRNGFGEVSGRLRADNTLVAVKSCRETLPPDL 530
QY 531 KAKFLQEARILKQYSHPNIVRLIGVCTOKOPITYIMVELVQGGDFLTFRTBGRRLKVKTL 590
DB 531 KAKFLQEARILKQYSHPNIVRLIGVCTOKOPITYIMVELVQGGDFLTFRTBGRRLKVKTL 590
QY 591 LQWVGDAAGMEYLESKCCIHRDLAARNCLVTEKNVLKISDFGMSREBAADGVYAAAGGLR 650
DB 591 LQWVGDAAGMEYLESKCCIHRDLAARNCLVTEKNVLKISDFGMSREBAADGVYAAAGGLR 650
QY 651 QVPEVKTAPBALNYGRYSSESVDVMSFGILLMETFSLGASPYPNLSNOQTRFEVKGGRLLP 710
DB 651 QVPEVKTAPBALNYGRYSSESVDVMSFGILLMETFSLGASPYPNLSNOQTRFEVKGGRLLP 710
QY 711 CPELCPDAVFRMLEQCMAYEPGQRPSEFSTTYOELQSIIRKRR 752
DB 711 CPELCPDAVFRMLEQCMAYEPGQRPSEFSTTYOELQSIIRKRR 752
QY 752 -----
DB 752 -----

```

```

TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
FILE REFERENCE: 50125/014002
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: US 60/222,081
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: DE 10030149.5
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 84
SOFTWARE: FaetsEQ for Windows Version 4.0
SEQ ID NO 64
LENGTH: 822
TYPE: prt
ORGANISM: Homo sapiens
US-09-886-319A-64

Query Match      47.8%; Score 1853; DB 4; Length 822;
Best Local Similarity 45.5%; Pred. No. 1e-120;
Matches 376; Conservative 150; Mismatches 217; Indels 84; Gaps 9;

QY 1 MGFSSSELCPGQGHVYLQOMQOAEELRLLEGMKMKMAQRYKSDREYAGLLHMSLQDSGGGS 60
DB 1 MGFSSSELCPGQGHVYLQOMQOAEELRLLEGMKMKMAQRYKSDREYAGLLHMSLQDSGGGS 60
QY 61 RAISPSPISSQSWAETTSQTEGSLRLRQHAEDLNSGPKSLSLIREQOLKRTYSEOW 120
DB 61 RAISPSPISSQSWAETTSQTEGSLRLRQHAEDLNSGPKSLSLIREQOLKRTYSEOW 120
QY 121 OQLOQELTKTHSODIKELKSYRALARDSAQAQKRYQEA KODKRDKADKRYSLMKLF 179
DB 121 OQLOQELTKTHSODIKELKSYRALARDSAQAQKRYQEA KODKRDKADKRYSLMKLF 179
QY 179 HMLHQQYVALAKGAOLHQQYDYDITPLPLDLSLQOQEMIKALNGIPDEYSQITSLYTE 238
DB 179 HMLHQQYVALAKGAOLHQQYDYDITPLPLDLSLQOQEMIKALNGIPDEYSQITSLYTE 238
QY 238 EYVAIHREMAAARLQPEAEYQGFILQYGSAPVPCVTFDESLLBEGEPLEBGELOQNE 299
DB 238 EYVAIHREMAAARLQPEAEYQGFILQYGSAPVPCVTFDESLLBEGEPLEBGELOQNE 299
QY 299 ELTVESVQHTLTSYTDLAVALATEMVFROEMVTQLOQELRNEBENTHPRERVOILLGRQYL 358
DB 299 ELTVESVQHTLTSYTDLAVALATEMVFROEMVTQLOQELRNEBENTHPRERVOILLGRQYL 358
QY 358 NLTAESLQWMLKTLAELMLQOQMLNKEAVLEBKRIEESSECEKSDIVLLLSQKQ 358
DB 358 NLTAESLQWMLKTLAELMLQOQMLNKEAVLEBKRIEESSECEKSDIVLLLSQKQ 358
QY 359 VLOEALQGLQVALCSQAQLQAQOELLQTKLEHLGEBEPVLLLODDRHSSTSSBOEREGGR 418
DB 359 VLOEALQGLQVALCSQAQLQAQOELLQTKLEHLGEBEPVLLLODDRHSSTSSBOEREGGR 418
QY 418 ALBELKOSVQOLRCEAKFSQKELLEQKVDNDGKEPPVNVYEDDASVYSMERKE-- 416
DB 418 ALBELKOSVQOLRCEAKFSQKELLEQKVDNDGKEPPVNVYEDDASVYSMERKE-- 416
QY 416 GRTPTLEILKSHISGIFR-PK----- 438
DB 416 GRTPTLEILKSHISGIFR-PK----- 438
QY 438 -----FSNLYRLGEGBPPSIPLIDHLLSTQ 464
DB 438 -----FSNLYRLGEGBPPSIPLIDHLLSTQ 464
QY 464 KQGDPLVESHGKPEEYVLSVSDGQRHFIQYVDNMYRFGGTGFSNIPQIIDHLYTTK 535
DB 464 KQGDPLVESHGKPEEYVLSVSDGQRHFIQYVDNMYRFGGTGFSNIPQIIDHLYTTK 535
QY 535 QPLTKSGSVTLHRAVPKDKMVLNHEBDLVLEQDIGRNGFGEVSGRLRADNTLVAVKSCR 523
DB 535 QPLTKSGSVTLHRAVPKDKMVLNHEBDLVLEQDIGRNGFGEVSGRLRADNTLVAVKSCR 523
QY 523 QVITKSGSVTLHRAVPKDKMVLNHEBDLVLEQDIGRNGFGEVSGRLRADNTLVAVKSCR 594
DB 523 QVITKSGSVTLHRAVPKDKMVLNHEBDLVLEQDIGRNGFGEVSGRLRADNTLVAVKSCR 594
QY 594 ETLPPDLKAKFLQEARILKQYSHPNIVRLIGVCTOKOPITYIMVELVQGGDFLTFRTBGA 583
DB 594 ETLPPDLKAKFLQEARILKQYSHPNIVRLIGVCTOKOPITYIMVELVQGGDFLTFRTBGA 583
QY 583 EDLPQELKTLQOEAELKQYSHPNIVRLIGVCTOKOPITYIMVELVQGGDFLTFRTBGA 654
DB 583 EDLPQELKTLQOEAELKQYSHPNIVRLIGVCTOKOPITYIMVELVQGGDFLTFRTBGA 654
QY 654 ELKQYVKSFLDAAGMLYLESKCCIHRDLAARNCLVTEKNVLKISDFGMSREBAADGVY 714
DB 654 ELKQYVKSFLDAAGMLYLESKCCIHRDLAARNCLVTEKNVLKISDFGMSREBAADGVY 714
QY 714 AASGGLQVYVKTAPBALNYGRYSSESVDVMSFGILLMETFSLGASPYPNLSNOQTRFEV 703
DB 714 AASGGLQVYVKTAPBALNYGRYSSESVDVMSFGILLMETFSLGASPYPNLSNOQTRFEV 703
QY 703 SSS-GIKQIPKMTAPBALNYGRYSSESVDVMSFGILLMETFSLGASPYPNLSNOQTRFEV 773
DB 703 SSS-GIKQIPKMTAPBALNYGRYSSESVDVMSFGILLMETFSLGASPYPNLSNOQTRFEV 773

```

```

RESULT 6
US-09-886-319A-64
; Sequence 64, Application US/09886319A
; Patent No. 6586185
; GENERAL INFORMATION:
; APPLICANT: Wolf, Eckard
; APPLICANT: Werner, Sabine
; APPLICANT: Halle, Jörn-Peter
; APPLICANT: Regenbogen, Johannes
; APPLICANT: Goppel, Andreas
; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for

```


QY	300	ELVNESVOTHTLSVTBELVATVEMVPRREWOTVLOQOELRYBDEENTHR-ERVOLLEKQ	358
QY	300	ELVNESVOTHTLSVTBELVATVEMVPRREWOTVLOQOELRYBDEENTHR-ERVOLLEKQ	358
Db	299	NULNADSLQVMTKLTAABELTOTQOMLHKRAVULBEKTI BBSFELCEKKSIVULLLOQK	356
QY	359	VLOEALQLOVALCSQAKLOAQOELLQTKLEHIGBERRPVULLLODDHSTSSSEOREG	418
Db	359	ALBELQSVQOALFCESEAKCAOAKALLEKQVQENDEKEPRPVVUNEEDARSVTSMERKE--	416
QY	419	GRFPTLEIKSHISGIFR-PK-----	438
Db	417	RLSKFEBSIRHSTINGITIKSPKSVLGSSTQVCVIVSGVRPLAENDHNGAIPRLAEODELL	475
QY	439	-----FSNLRYLGEGBFSPSRLIDHLLST	463
Db	476	KQOGDELVRBESHGKPEBEVULSVYSDGQRHFIIOQVUDLYLPRFEGTGFNSINPOLIDHHFNT	533
QY	464	QOQLTKKSGVULHRAVRXD-KWVLLNHEDI-VLGEQVGRNPFGBVRGRLRANLTVAVXSC	522
Db	536	KQVITTKSGVULNLPFKDKVLLNHEDVLSBELJGKNGFVGVLTKL-KXTVAALTKC	594
QY	523	RETLPRDLKAKFLQOEARILKOYSHPNVYRLIGVCTOKOPIYIVMELVQSGDLFTPLRTG	582
Db	595	KEDLPQGLKTFQOEARILKOYDHPNIVYGLIGVCTORQVYIIMELVGGGDLFTPLRARK	654
QY	583	ARLRVYTLLOMVDAAAGMEYLESKKCHRDILAANCIVTEBNTLXISDPFMSFEEDAV	642
Db	655	DELKTLKQVRFSLDVAAGMUYLESKCHIRDLAARNCIVAGENNTLXISDPFMSIQBOGGV	714
QY	643	YAAAGGLRQVFPVMTAPALNATGRYSSESVDVMSFGILLMETFSLGASFPYPLINSQOQTRF	702
Db	715	YSSS-GIKQIPKMTAPALNATGRYSSESVDVMSFGILLMETFSLGVCYCPGMATQOAREQ	773
QY	703	VEBGGRLPCBCLCPDAVFLMEOCMAYEGQORPSSTIYQOELQISRK	749
Db	774	VEBGRYMSAPQNCPEEVLFTIMKCMWYKKNPKKPNNDLKLTVYIKK	820

RESULT 9
 US-08 701-191A-38
 Sequence 38, Application US/08701191A
 Patent No. 5942428
 GENERAL INFORMATION:
 APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
 APPLICANT: and Steven R. Hubbard
 TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 633 West Filth Street
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 MEDIUM TYPE: storage
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: IBM P.C. DOS 5.0
 SOFTWARE: Fastseq for Windows 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/701,191A
 FILING DATE: August 21, 1996
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Walburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 32227/088

```

? TELECOMMUNICATION INFORMATION
? TELEPHONE: (213) 489-1600
? TELEFAX: (213) 955-0440
?
? TELEX: 67-3510
?
? INFORMATION FOR SEQ ID NO: 38:
?
? SEQUENCE CHARACTERISTICS:
? LENGTH: 275 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
?
? MOLECULE TYPE: protein
?
? OS-08-701-191A-38

```

Query Match	37.2%	Score	1444	DB 2	Length	275	
Best Local Similarity	99.6%	Pred. NO.	7.4e-93				
Matches 274	0	Mismatches	1	Indels	0	Gaps	0

Qy	478	AVPKDQKVLNHNHDLVYGEIOGRNGEYVFSRLADNTLVAVKSCRETLPPDLAKFLOE	537
Db	1	AVPKDKVLNHNHDLVYGEIOGRNGEYVFSRLADNTLVAVKSCRETLPPDLAKFLOE	60
Qy	538	ARILKOYSHPNIVALLIGVCTOKOPYIYIMELVOGDFLTFRTEGARLARVKTLLQWYGA	537
Db	61	ARILKOYSHPNIVRLIGVCTOKOPYIYIMELVOGDFLTFRTEGARLARVKTLLQWYGA	120
Qy	598	AAGMEYLEKSCITTHDLAARNCIVYEKQVLYISDPGMSREEDSVYVAASGGLROVPVKT	657
Db	121	AAGMEYLEKSCITTHDLAARNCIVYEKQVLYISDPGMSREEDSVYVAASGSGNQVPVKT	180
Qy	658	APBALNYGYSSESQVMSFGILLMTFSLGASPPYPNLSNOOTREFVEKGRLLPCPELCPD	717
Db	181	APBALNYGYSSESQVMSFGILLMTFSLGASPPYPNLSNOOTREFVEKGRLLPCPELCPD	240
Qy	718	AVFRLMEQCMAYEGQRPSEFTIYVELOSIKRHR	752
Db	241	AVFRLMEQCMAYEGQRPSEFTIYVELOSIKRHR	275

RESULT 10
 US-09-664-526-38
 Sequence 38, Application US/09664526
 Patent No. 6682921
 GENERAL INFORMATION:
 APPLICANT: MOHAMMADI, MOOSA
 APPLICANT: SCHLESINGER, JOSEPH
 APPLICANT: HUBBARD, STEVAN R.
 TITLE OF INVENTION: RECEPTOR TYROSINE KINASES
 TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN OF NON-INSULIN
 FILE REFERENCE: 038602/0847
 CURRENT APPLICATION NUMBER: US/09/664,526
 CURRENT FILING DATE: 2000-09-18
 PRIOR APPLICATION NUMBER: 09/188, 809
 PRIOR FILING DATE: 1998-11-09
 PRIOR APPLICATION NUMBER: 08/701,191
 PRIOR FILING DATE: 1996-08-21
 NUMBER OF SEQ ID NOS: 41
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 38
 LENGTH: 275
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-664-526-38

	Query Match	37.2%	Score 1444	DB 4	Length 275
	Best Local Similarity	99.6%	Pred. No. 7.4e-93		
	Matches 274	Conservative	0	Mismatches 1	Indels 0
					Gaps 0
Qy	478	AVPKDKVNLNEDLVLEGEIGRGNGFGEVFSRRLPADNTLVAVKSCRETLPPDLKAKFLQE	537		
Db	1	AVPKDKVNLNEDLVLEGEIGRGNGFGEVFSRRLPADNTLVAVKSCRETLPPDLKAKFLQE	60		
Qy	538	ARILKQYSHENIVRLIGVCTOKQDIYYVMEVQGGDFLTFLPTBSARLRVKTLLQWNGDA	597		

Db 61 ARLKQSHPNIVRLIGVCTQKQPIYIWMELVGGDFLFLRTGRLVKTLQWVGA 120
Qy 558 AAGMEYLSKCCIRHDLAARNCLVTEKNVLTISDFGMSREEDGVYAASGGLRQVPVKMT 657
Db 121 AAGMEYLSKCCIRHDLAARNCLVTEKNVLTISDFGMSREEDGVYAASGGSROVPVKMT 180
Qy 658 APEALNYGVSSESDVWSFGILLMETFSLGASPPYPLNSNOQTRFPEKGRLLPCPELCPD 717
Db 121 APEALNYGVSSESDVWSFGILLMETFSLGASPPYPLNSNOQTRFPEKGRLLPCPELCPD 240
Qy 718 AVFRLMEQCMAYEPGGRPSFTTYOELQSIKRRH 752
Db 241 AVFRLMEQCMAYEPGGRPSFTTYOELQSIKRRH 275

RESULT 11
US-07-857-224B-62
; Sequence 62, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (international) 41 1 632 2830
; TELEFAX: (international) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; ORIGINAL SOURCE:
; ORGANISM: human
; FEATURE: Protein kinase; Table 8 Column 70
; PUBLICATION INFORMATION:
; AUTHORS:
; AUTHORS: Hanke, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988
; US-07-857-224B-62

Query Match 35.2%; Score 1364.5; DB 2; Length 262;
Best Local Similarity 99.6%; Pred. No. 2.4e-87;
Matches 262; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 489 EDVLVGEQIGRNGFVSGRLADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 548
Db 1 EDVLVGEQIGRNGFVSGRLADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 60

Qy 549 IYRLIGVCTQKQPIYIWMELVGGDFLFLRTGRLVKTLQWVGAAGMEYLSKC 608
Db 61 IYRLIGVCTQKQPIYIWMELVGGDFLFLRTGRLVKTLQWVGAAGMEYLSKC 120
Qy 609 CIHRDLAARNCLVTEKNVLTISDFGMSREEDGVYAASGGLRQVPVKMTAPEALNYGRYS 668
Db 121 CIHRDLAARNCLVTEKNVLTISDFGMSREEDGVYAASGGLRQVPVKMTAPEALNYGRYS 179
Qy 669 SSSDVWSFGILLMETFSLGASPPYPLNSNOQTRFPEKGRLLPCPELCPDAVFLMEQCA 728
Db 180 SSSDVWSFGILLMETFSLGASPPYPLNSNOQTRFPEKGRLLPCPELCPDAVFLMEQCA 239
Qy 729 YEPGGRPSFTTYOELQSIKRRH 751
Db 240 YEPGGRPSFTTYOELQSIKRRH 262

RESULT 12
US-07-857-224B-63
; Sequence 63, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) CH-8092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/857,224B
; FILING DATE: 03/25/92
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA: none
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (international) 41 1 632 2830
; TELEFAX: (international) 41 1 262 2437
; TELEX: none
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; ORIGINAL SOURCE:
; ORGANISM: cat
; FEATURE: Protein kinase; Table 8 Column 71
; PUBLICATION INFORMATION:
; AUTHORS:
; AUTHORS: Hanke, S. K.
; AUTHORS: Quinn, A. M.
; AUTHORS: Hunter, T.
; TITLE: The protein kinase family
; JOURNAL: Science
; VOLUME: 241
; PAGES: 42-52
; DATE: 1988
; US-07-857-224B-63

Query Match 34.8%; Score 1350.5; DB 2; Length 262;
Best Local Similarity 97.7%; Pred. No. 2.2e-86;
Matches 257; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 489 EDLVIGEIOGRNFGVEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHN 548
DB 1 EDLVIGEIOGRNFGVEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHN 60
QY 549 IYRLIGVCTOKOPITYIMELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 608
DB 61 IYRLIGVCTOKOPITYIMELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 120
QY 609 CIHRDLAARNCLVTEKNVLTISDFGMSREADVGYAASGGLRQVYKWTAPALNYGRYS 668
DB 121 CIHRDLAARNCLVTEKNVLTISDFGMSREADVGYAASGGLRQVYKWTAPALNYGRYS 179
QY 669 SSSDVMSFGILLMEFSLGASPYPLNSNOOTREFVEKGRPLCPBLCPDAVRLMEQCA 728
DB 180 SSSDVMSFGILLMEFSLGASPYPLNSNOOTREFVEKGRPLCPBLCPDAVRLMEQCA 239
QY 729 YEPGQRPFSFTTYOELQSIKRRH 751
DB 240 YEPGQRPFSFTTYOELQSIKRRH 262

RESULT 13

US-07-857-224B-64
Sequence 64, Application US/07857224B
Patent No. 5958784

GENERAL INFORMATION:

APPLICANT: Benner, Steven A.
TITLE OF INVENTION: Predicting Folded Structures of Proteins
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven A. Benner
STREET: Hadlaubstrasse 151
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8092

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage

COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0

SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/857,224B

CLASSIFICATION: 436
FILING DATE: 03/25/92

PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (international) 41 1 632 2830
TELEFAX: (international) 41 1 262 2437

TELEX: none
INFORMATION FOR SEQ ID NO: 64:

SEQUENCE CHARACTERISTICS:
LENGTH: 262

TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE:
DESCRIPTION: protein

ORGANISM: chicken
FEATURE: Protein kinase, Table 8 Column 72
PUBLICATION INFORMATION:

AUTHORS:
AUTHORS: Hanko, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.

TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988

US-07-857-224B-64

Query Match

28.8%; Score 1116.5; DB 2; Length 262;

Best Local Similarity 79.8%; Pred. No. 4.2e-70;
Matches 210; Conservative 29; Mismatches 23; Indels 1; Gaps 1;
QY 489 EDLVIGEIOGRNFGVEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHN 548
DB 1 EDLVIGEIOGRNFGVEVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOYSHN 60
QY 549 IYRLIGVCTOKOPITYIMELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 608
DB 61 IYRLIGVCTOKOPITYIMELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 120
QY 609 CIHRDLAARNCLVTEKNVLTISDFGMSREADVGYAASGGLRQVYKWTAPALNYGRYS 668
DB 121 CIHRDLAARNCLVTEKNVLTISDFGMSREADVGYAASGGLRQVYKWTAPALNYGRYS 179
QY 669 SSSDVMSFGILLMEFSLGASPYPLNSNOOTREFVEKGRPLCPBLCPDAVRLMEQCA 728
DB 180 SSSDVMSFGILLMEFSLGASPYPLNSNOOTREFVEKGRPLCPBLCPDAVRLMEQCA 239
QY 729 YEPGQRPFSFTTYOELQSIKRRH 751
DB 240 YEPGQRPFSFTTYOELQSIKRRH 262

RESULT 14

US-10-003-690-8
Sequence 8, Application US/10003690
Patent No. 6787345

GENERAL INFORMATION:

APPLICANT: Roy A.J. Curtis
TITLE OF INVENTION: 55053, A No. 6787345e1 Human Eukaryotic Kinase
TITLE OF INVENTION: and Uses Therefor
FILE REFERENCE: MNI-206
CURRENT APPLICATION NUMBER: US/10/003,690
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/248,893
PRIOR FILING DATE: 2000-11-15
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8

LENGTH: 280
TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: tyrkin_6 domain

US-10-003-690-8

Query Match 16.6%; Score 642; DB 4; Length 280;
Best Local Similarity 46.3%; Pred. No. 4.7e-37;
Matches 133; Conservative 49; Mismatches 65; Indels 40; Gaps 8;

QY 491 IYLGEOIGRNGFVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOY--SHN 548
DB 1 IYLGEOIGRNGFVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKOY--SHN 53
QY 549 IYRLIGVCTOK-----QPIYIMELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 601
DB 54 IYRLIGVCTOK-----QPIYIMELVGGDFLTLRTGRLRVKTLLOWGDAAGMEYLESKC 113
QY 602 EYLESKCCIHRLAARNCLVTEKNVLTISDFGMSRE-----EADGYV---AASGLR 650
DB 114 EYLESKCCIHRLAARNCLVTEKNVLTISDFGMSRE-----EADGYV---AASGLR 173
QY 651 QYVPMWTAPEALNYGRYSSESDVMSFGILLMEFSLGASPYPLNSNOOTREFVEK 705
DB 174 QYVPMWTAPEALNYGRYSSESDVMSFGILLMEFSLGASPYPLNSNOOTREFVEK 233
QY 706 GGLRCPBLCPDAVRLMEQCAVYKWTAPALNYGRYSSESDVMSFGILLMEFSLGASPYPLNSNOOTREFVEK 744
DB 234 GGLRCPBLCPDAVRLMEQCAVYKWTAPALNYGRYSSESDVMSFGILLMEFSLGASPYPLNSNOOTREFVEK 280

RESULT 15

```
US-09-949-016-6798
; Sequence 6798, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6798
; LENGTH: 1148
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6798

Query Match      16.1%; Score 624.5; DB 4; Length 1148;
Beet Local Similarity 36.0%; Pred. No. 4.5e-35;
Matches 138; Conservative 68; Mismatches 128; Indels 49; Gaps 10;

QY 405 DNHSTSSSEQREREGRTPTLEILKSHISGIFRPKFS----- 440
DB 141 EKHSHVHGVSNAAEYF-----LSGINSFLVRESESSPSQSTSLRYEGRVYHRYNT 196
QY 441 ----NLYRLEGGFPEIPLIDH-----LIST--QOPLTKSGVVLHRAVPK-DKVTLN 487
DB 197 ASDGKLTVSSESRFNTLAEVLVHHSTVADGLITTLHYPAKRNKPTVGVSPNYDKEME 256
QY 488 HEDVLVBGIGRNGEVEFSGRLDNTLVAVKSCRE-TLPDLYAKKFLQEARILKQYSH 546
DB 257 RTDITMKHKLGGQGEVYEGVWKYSLTVAVKTKLEDTEWE--EFLKEAIVMKEIRH 313
QY 547 PNIVRLIGVCTOKOPIYIVMELVGGDFLPLRTGEGARLRVK--TLLOKVGDAAGMEYL 604
DB 314 PVLVQLLGVCRTREPPYIITTEMTYGNLDYLR-ECNROEVNAVVLVMAIQISSAMEYL 372
QY 605 ESKCCIHRDLAARNCLVTEKNVLIKISDFGMSRBEADGVYASGGLRQVPVKYTAPBALNY 664
DB 373 EKNRTHRDLAARNCLVGENHLVKVADFGLSRLMTGDTYTAHAGAK-FPIKWTAPESLAY 431
QY 665 GRYSSESVDVMSFGILLMETFSLGASPYPNLSNOQTRFEVEKGRLLPCPELCPDAVFRIME 724
DB 432 NKFISIKSDVMAFGVLMELATYGMSPYPIDRSQVYELLEKDYRMKRPBGCEPKYELMR 491
QY 725 QCMAYEPGQPSFSSTIYOELOSI 747
DB 492 ACMQWNPSPDRPSFAELHQAFTM 514
```

Search completed: March 18, 2005, 16:01:08
Job time : 57 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 15:50:13 ; Search time 45 Seconds

(without alignments)
1607.889 Million cell updates/sec

Title: US-10-660-763-2

Perfect score: 3879

Sequence: 1 MGFSSELCSPOGHGVLQGMQ.....QRPSTTYQLSIRKXNR 752

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

283416

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Database :

1: PIR 79:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3834	98.8	822	1	TVHUF	protein-tyrosine k
2	3620	93.3	820	1	TVCTFF	protein-tyrosine k
3	3473	89.5	820	1	I48347	protein-tyrosine k
4	2635	67.9	824	2	I50618	c-fes proto onco
5	2570	66.3	823	1	TVFVFS	protein-tyrosine k
6	2561	66.0	873	1	TVFVFF	protein-tyrosine k
7	2479	63.9	609	1	TVMVC	protein-tyrosine k
8	1960	51.0	477	1	TVMVC	protein-tyrosine k
9	1853	47.8	822	1	TVHUF	protein-tyrosine k
10	1450	37.4	533	1	TVFVFF	protein-tyrosine k
11	1160	29.9	803	1	OKPFS	protein-tyrosine k
12	1099.5	28.3	323	2	S04328	protein-tyrosine k
13	1097	28.3	453	1	I49663	protein-tyrosine k
14	734	18.9	181	2	I50406	proto-fps protein
15	632	16.3	697	2	A26132	gag-abl-pol polyp
16	623	16.1	1123	2	A39962	kinase-related tra
17	622	16.0	1130	1	TVHUA	protein-tyrosine k
18	620.5	16.0	981	1	FOVMGM	gag-abl polypotei
19	619.5	16.0	334	2	S24552	protein-tyrosine k
20	608	15.7	505	2	I59296	protein-tyrosine k
21	605.5	15.6	1146	2	B35962	protein-tyrosine k
22	605.5	15.6	1182	2	A35962	protein-tyrosine k
23	604	15.6	465	2	I48926	protein-tyrosine k
24	603.5	15.6	527	2	I49133	protein-tyrosine k
25	603.5	15.6	1196	2	T23832	protein-tyrosine k
26	603.5	15.6	1520	1	TVFPA	protein-tyrosine k
27	602	15.5	507	2	A55625	protein-tyrosine k
28	600.5	15.5	467	2	I56579	protein-tyrosine k
29	597	15.4	527	2	S13763	protein-tyrosine k

30	595.5	15.4	362	2	S24551	protein-tyrosine k
31	592.5	15.3	521	2	I56997	protein-tyrosine k
32	591	15.2	637	2	I84483	protein-tyrosine k
33	589	15.2	675	2	S60612	protein-tyrosine k
34	588.5	15.2	620	1	S33253	protein-tyrosine k
35	586.5	15.1	1028	2	A45388	protein-tyrosine k
36	583	15.0	507	1	A39939	protein-tyrosine k
37	583	15.0	2340	2	I48310	kinase-related pro
38	581.5	15.0	680	4	I38491	nucleophosmin/ana
39	581	15.0	630	1	T01380	protein-tyrosine k
40	580.5	15.0	496	2	A56040	protein-tyrosine k
41	579.5	14.9	402	1	TVFVUR	protein-tyrosine k
42	579.5	14.9	505	1	S24550	protein-tyrosine k
43	579	14.9	2347	1	TVHURS	kinase-related pro
44	578.5	14.9	590	1	TVFVFS	protein-tyrosine k
45	578	14.9	1068	2	I51670	focal adhesion kin

ALIGNMENTS

```
RESULT 1
TVHUF
protein-tyrosine kinase (EC 2.7.1.112) fes/fps - human
C.Species: Homo sapiens (man)
C.Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #ext_change 03-Jul-2004
C.Accession: A24673; A60188
R.Roebroek, A.J.M.; Schalken, J.A.; Verbeek, J.S.; Van den Ouweland, A.M.W.; Onnekink, C.
EMBO J. 4, 2897-2903, 1985
A.Title: The structure of the human c-fes/fps proto-oncogene.
A.Reference number: A24673; MUID:86055727; PMID:4065096
A.Accession: A24673
A.Molecule type: DNA
A.Residues: 1-822 <ROE>
A.Cross-references: UNIPROT:P07332; GB:X06292; GB:M14209; GB:M14589; NID:G31348; PIDN:CA7
R.Alicia, M.; Antolini, F.; Van de Ven, W.J.; Lanfrancone, L.; Grignani, F.; Pellicci, P.
Oncogene 5, 267-275, 1990
A.Title: Characterization of human and mouse c-fes cDNA clones and identification of the
A.Reference number: A60188; MUID:90191711; PMID:2179616
A.Accession: A60188
A.Molecule type: mRNA
A.Residues: 1-718, 'S', 720-822 <ALC>
A.Cross-references: GB:X52192; NID:929890; PIDN:CA36438.1; PID:G29891
A.Note: the authors translated the codon TCC for residue 719 as Leu
C.genetics:
A.Gene: GDB:FES
A.Cross-references: GDB:119906; OMIM:190030
A.Map position: 15q26.1-15q26.1
A.Introns: 71/3; 129/3; 162/1; 223/2; 269/2; 309/2; 412/3; 440/3; 510/3; 51/3; 56/3
C.Function:
A.Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C.Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C.Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phos
P.2-822/Product: protein-tyrosine kinase fes/fps #status predicted <MTR>
P.460-545/Domain: SH2 homology <SH2>
P.559-821/Domain: protein kinase homology <KIN>
P.567-575/Region: protein kinase ATP-binding motif
P.2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
P.590/Active site: Lys #status predicted

Query Match      98.8%  Score 3834;  DB 1;  Length 822;
Best Local Similarity 91.5%;  Pred. No. 7.8e-141;
Matches 752;  Conservative 0;  Mismatches 0;  Indels 70;  Gaps 1;
```

```

Db      121 OOLQOELTKTHSODIEKLSQSYRALARDSQAQKRYQASKODKRDKADKVRSLMKLF 180
Qy      181 AHNHRYVLGVRAAQLHHQHHLHQLLPGLRLSLQDLHHEBACTIKELIQEYLSISLVODE 240
Db      181 AHNHRYVLGVRAAQLHHQHHLHQLLPGLRLSLQDLHHEBACTIKELIQEYLSISLVODE 240
Qy      241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPDVPFCVTFDESLEEGEPLEPGELQANE 300
Db      241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPDVPFCVTFDESLEEGEPLEPGELQANE 300
Qy      301 LTVESVOHTLTSVTDELAVATEMVFRRQEWVTQLQOELNBEENTHPRERVOGLLGRQYVL 360
Db      301 LTVESVOHTLTSVTDELAVATEMVFRRQEWVTQLQOELNBEENTHPRERVOGLLGRQYVL 360
Qy      361 QEALQGLQVALCSQAKLQAOQELLOTKEHLGEGEPPLVLLQDDRHSTSSSQEREGGR 420
Db      361 QEALQGLQVALCSQAKLQAOQELLOTKEHLGEGEPPLVLLQDDRHSTSSSQEREGGR 420
Qy      421 TPTLEILKSHISGIRPKFS----- 440
Db      421 TPTLEILKSHISGIRPKFS----- 440
Qy      441 -----NLYRLEGEGEPSPILIDHLSTQOPLTK 470
Db      481 LVRESQKQEVYLVLMQGLPRHFIQSIDNLYRLEGEGEPSPILIDHLSTQOPLTK 540
Qy      471 SGVYLHRAVPKQKWLNHEDVLAGEQIGRNGEVSGLRADNTLVAVKSCRETLPPDL 530
Db      541 SGVYLHRAVPKQKWLNHEDVLAGEQIGRNGEVSGLRADNTLVAVKSCRETLPPDL 600
Qy      531 KAKFLOEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFRTGEGARLRVKT 590
Db      601 KAKFLOEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFRTGEGARLRVKT 660
Qy      591 LQWVGDAAGMEYLESKCCIHRLDIAARNCIVTEKNVLSDFGMSREEDGVYAASGGLR 650
Db      661 LQWVGDAAGMEYLESKCCIHRLDIAARNCIVTEKNVLSDFGMSREEDGVYAASGGLR 720
Qy      651 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSIGASPYPNLSNQOTREFVEKGRPL 710
Db      721 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSIGASPYPNLSNQOTREFVEKGRPL 780
Qy      711 CPELCDAVFRLEMCQWAVEPQGRPSFSSTIYQELQSIKRRHR 752
Db      781 CPELCDAVFRLEMCQWAVEPQGRPSFSSTIYQELQSIKRRHR 822

```

RESULT 2

TVCTFF

protein-tyrosine kinase (BC 2.7.1.112) fes/fps - cat

C/Species: Felis silvestris catus (domestic cat)

C/Date: 30-Sep-1989 #sequence_revision 30-Jun-1992 #ext_change 23-Feb-1997

C/Accession: A27824

R./Author: A.J.M.; Schalken, J.A.; Omekink, C.; Bloemers, H.P.J.; Van de Ven, W.J.M.

J. Virol. 61, 2009-2016, 1987

A./Title: Structure of the feline c-fes/fps proto-oncogene: genesis of a retroviral oncog

A./Reference number: A27824; PMID:87198954; PMID:3553615

A./Accession: A27824

A./Molecule type: DNA

A./Residues: 1-820 <ROE>

C/Genetics:

A./Gene: fes/fps

C./Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology

C./Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; proto-oncogene

F/458-543/Domain: SH2 homology <SH2>

F/557-819/Domain: protein kinase homology <KIN>

F/565-573/Region: protein kinase ATP-binding motif

F/588/Active site: Lys #status predicted

F/711/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 93.3%; Score 3620; DB 1; Length 820;

Best Local Similarity 85.8%; Pred. No. 1.3e-132;

```

Matches 705; Conservative 21; Mismatches 24; Indels 72; Gaps 2;
Qy      1 MGFSEELSPQGHVLAQMOAEALRLLEGMRKMAQRYASDEEYAGLLHHMSLDQSGGS 60
Db      1 MGFSEELSPQGHVLAQMOAEALRLLEGMRKMAQRYASDEEYAGLLHHMSLDQSGG-- 58
Qy      61 RAISPDSPISQSWAEITSQTEGLSRLLRQHAEDLNSGSLSLIRERQURKTYSEGM 120
Db      59 RGTGYSPISQSWAEITSQTEGLSRLLRQHAEDLNSGSLSLIRERQURKTYSEGM 118
Qy      121 OOLQOELTKTHSODIEKLSQSYRALARDSQAQKRYQASKODKRDKADKVRSLMKLF 180
Db      119 OOLQOELTKTHSODIEKLSQSYRALARDSQAQKRYQASKODKRDKADKVRSLMKLF 178
Qy      181 AHNHRYVLGVRAAQLHHQHHLHQLLPGLRLSLQDLHHEBACTIKELIQEYLSISLVODE 240
Db      179 AHNHRYVLGVRAAQLHHQHHLHQLLPGLRLSLQDLHHEBACTIKELIQEYLSISLVODE 238
Qy      241 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPDVPFCVTFDESLEEGEPLEPGELQANE 300
Db      239 VVAIHREMAAAARIQPEAEYQGFRLQYGSAPDVPFCVTFDESLEEGEPLEPGELQANE 298
Qy      301 LTVESVOHTLTSVTDELAVATEMVFRRQEWVTQLQOELNBEENTHPRERVOGLLGRQYVL 360
Db      299 LTVESVOHTLTSVTDELAVATEMVFRRQEWVTQLQOELNBEENTHPRERVOGLLGRQYVL 358
Qy      361 QEALQGLQVALCSQAKLQAOQELLOTKEHLGEGEPPLVLLQDDRHSTSSSQEREGGR 420
Db      359 QEALQGLQVALCSQAKLQAOQELLOTKEHLGEGEPPLVLLQDDRHSTSSSQEREGGR 418
Qy      421 TPTLEILKSHISGIRPKFS----- 440
Db      419 TPTLEILKSHISGIRPKFS----- 440
Qy      441 -----NLYRLEGEGEPSPILIDHLSTQOPLTK 470
Db      479 LVRESQKQEVYLVLMQGLPRHFIQSADNLYRLEGEGEPSPILIDHLSTQOPLTK 538
Qy      471 SGVYLHRAVPKQKWLNHEDVLAGEQIGRNGEVSGLRADNTLVAVKSCRETLPPDL 530
Db      539 SGVYLHRAVPKQKWLNHEDVLAGEQIGRNGEVSGLRADNTLVAVKSCRETLPPDL 598
Qy      531 KAKFLOEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFRTGEGARLRVKT 590
Db      599 KAKFLOEARILKQYSHPNIVRLIGVCTQKOPRIYIMELVQGGDFLTFRTGEGARLRVKT 658
Qy      591 LQWVGDAAGMEYLESKCCIHRLDIAARNCIVTEKNVLSDFGMSREEDGVYAASGGLR 650
Db      659 LQWVGDAAGMEYLESKCCIHRLDIAARNCIVTEKNVLSDFGMSREEDGVYAASGGLR 718
Qy      651 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSIGASPYPNLSNQOTREFVEKGRPL 710
Db      719 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSIGASPYPNLSNQOTREFVEKGRPL 778
Qy      711 CPELCDAVFRLEMCQWAVEPQGRPSFSSTIYQELQSIKRRHR 752
Db      779 CPELCDAVFRLEMCQWAVEPQGRPSFSSTIYQELQSIKRRHR 820

```

RESULT 3

I48347

protein-tyrosine kinase-related protein - mouse

N./Alternate names: c-fes protein

C./Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #ext_change 09-Jul-2004

C/Accession: I48347; J01112

R./Author: A.F.; Kurban, R.R.

Oncogene 3, 289-294, 1988

A./Title: Isolation and structural analysis of murine c-fes cDNA clones.

A./Reference number: I48347; PMID:89083198; PMID:3060793

A./Accession: I48347

A./Status: preliminary; translated from GB/EMBL/DBJ

A./Molecule type: mRNA

A:Residues: 1-820 <RES>
A:Cross-references: UNIPROT:P16879; EMBL:X12616; NID:g50955; PIDD:CAA1138.1; PID:g50956
R:Wills, A.F.; Kurben, R.R.; Hovens, C.M.; Ralph, S.U.
Gene 85, 67-74, 1989
A:Title: The application of the polymerase chain reaction to cloning members of the pro
A:Reference number: JH0112; MUID:90152381; PMID:2482828
A:Accession: JH0112
A:Molecule type: mRNA
A:Residues: 678-745 <WIL>
A:Experimental source: haemopoietic cell, clone FD15
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP
F:458-543/Domain: SH2 homology <SH2>
F:557-819/Domain: protein kinase homology <KIN>
F:565-573/Region: protein kinase ATP-binding motif

Query Match 89.5%; Score 3473; DB 2; Length 820;
Best Local Similarity 82.4%; Pred. No. 66-127;
Matches 677; Conservative 35; Mismatches 38; Indels 72; Gaps 2;

QY 1 MGFSSSLCPQGHGVLQOQOEALRLLEGKRMMAORVYSDBRYAGLHHMSIQDSGQS 60
DB 1 MGFSSSLCPQGHGVLQOQOEALRLLEGKRMMAORVYSDBRYAGLHHMSIQDSGQS 60
QY 61 RAISPSPISSQSAEITTSOTEGSLRLRQHAEDLNSGPIKSLILRERQOLRKTYSBQM 120
DB 61 WSGSPSPSSQSAEITTSOTEGSLRLRQHAEDLNSGPIKSLILRERQOLRKTYSBQM 120
QY 121 QOLQOELTHTSODIEKLKSQYRALARDASQAKRYQASKDKDRKADKXVRSIMKLF 180
DB 121 QOLQOELTHTSODIEKLKSQYRALARDASQAKRYQASKDKDRKADKXVRSIMKLF 180
QY 181 AHNHRYVLGVRAQLHGHQHQLLPGLRLSLQDLHEEMACILKEILQETLSSIVQD 240
DB 181 AHNHRYVLGVRAQLHGHQHQLLPGLRLSLQDLHEEMACILKEILQETLSSIVQD 240
QY 241 VVAIHHEMAAARIQPEAEYQGLROYGSAPRVPCTPFDESILREGEPLREGEOLNE 300
DB 241 VVAIHHEMAAARIQPEAEYQGLROYGSAPRVPCTPFDESILREGEPLREGEOLNE 300
QY 301 LTVESVOHTLTSVTDLAVALATEMVPFRQEMVTOLOQELRNEBENTHPRERVOQLGROYL 360
DB 301 LTVESVOHTLTSVTDLAVALATEMVPFRQEMVTOLOQELRNEBENTHPRERVOQLGROYL 360
QY 361 QEALQGLQVALCSQAKLQAOQELLOTKLEHLGEGEPVLLIQQDRHSTSSSEQER 420
DB 361 QEALQGLQVALCSQAKLQAOQELLOTKLEHLGEGEPVLLIQQDRHSTSSSEQER 420
QY 421 TPTLEILKSHISGIRPKFS----- 440
DB 421 TPTLEILKSHISGIRPKFS----- 440
QY 441 -----NLVRLGEGEPSPILPLIDHLSTQOPLTKK 470
DB 441 -----NLVRLGEGEPSPILPLIDHLSTQOPLTKK 470
QY 479 LVRESQKQEVYLVVMMDGHRHFIIOQLDNLVRLGEGEPSPILPLIDHLSTQOPLTKK 538
DB 479 LVRESQKQEVYLVVMMDGHRHFIIOQLDNLVRLGEGEPSPILPLIDHLSTQOPLTKK 538
QY 471 SGVILHRAVPKDKWLVNHEEDLVLSGQIGRNGEVSGLRLADNTLVAVKSCRETLPPDL 530
DB 471 SGVILHRAVPKDKWLVNHEEDLVLSGQIGRNGEVSGLRLADNTLVAVKSCRETLPPDL 530
QY 531 KAKPELOEALILKQYSHPNIVRLIGVCTQOPRIYVNEVLVOGGPILFPLRTGEGARLVKTL 590
DB 531 KAKPELOEALILKQYSHPNIVRLIGVCTQOPRIYVNEVLVOGGPILFPLRTGEGARLVKTL 590
QY 591 LQWVGDAAGMEYLESKCCHIRDLAARNCLVTEKNVLTISDFGMSREEDAGVAVASGAR 650
DB 591 LQWVGDAAGMEYLESKCCHIRDLAARNCLVTEKNVLTISDFGMSREEDAGVAVASGAR 650
QY 651 QVAVKATAPBALNYGYSSSDVWSFGILLMETFSLGASPYPNLSNQOTREFVEKGRLP 710
DB 651 QVAVKATAPBALNYGYSSSDVWSFGILLMETFSLGASPYPNLSNQOTREFVEKGRLP 710
QY 719 QVAVKATAPBALNYGYSSSDVWSFGILLMETFSLGASPYPNLTNQOTREFVEKGRLP 778
DB 719 QVAVKATAPBALNYGYSSSDVWSFGILLMETFSLGASPYPNLTNQOTREFVEKGRLP 778
QY 711 CPELCPDAVRLMEQCMAYEPGQRSPFSITICQELHSIRKRR 752
DB 711 CPELCPDAVRLMEQCMAYEPGQRSPFSITICQELHSIRKRR 752

DB 779 CPELCPDAVRLMEQCMAYEPGQRSPFSITICQELHSIRKRR 820
RESULT 4
150618
C:fps proto oncogene - chicken
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: 150618
R:Huang, C.C.; Hammond, C.; Bishop, J.M.
J. Mol. Biol. 181, 175-186, 1985
A:Title: Nucleotide sequence and topography of chicken c-fps. Genesis of a retroviral onc
A:Reference number: 150618; MUID:85160839; PMID:3879969
A:Accession: 150618
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-824 <HUA>
A:Cross-references: UNIPROT:Q90778; EMBL:X02266; NID:g63203; PIDD:CAA26155.1; PID:g87104;
C:genetics:
A:gene: c-fps
A:introns: 75/3; 133/3; 166/1; 227/2; 273/2; 313/2; 354/2; 416/3; 442/3; 512/3; 553/3; 5;
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP
F:462-547/Domain: SH2 homology <SH2>
F:561-823/Domain: protein kinase homology <KIN>
F:569-577/Region: protein kinase ATP-binding motif

Query Match 67.9%; Score 2635; DB 2; Length 824;
Best Local Similarity 62.7%; Pred. No. 1,1e-94;
Matches 518; Conservative 99; Mismatches 133; Indels 76; Gaps 3;

QY 1 MGFSSSLCPQGHGVLQOQOEALRLLEGKRMMAORVYSDBRYAGLHHM---SLQDS 56
DB 1 MGFSSSLCPQGHGVLQOQOEALRLLEGKRMMAORVYSDBRYAGLHHM---SLQDS 56
QY 57 GGOSPAISDPSISQSAEITTSOTEGSLRLRQHAEDLNSGPIKSLILRERQOLRKTY 116
DB 57 GGOSPAISDPSISQSAEITTSOTEGSLRLRQHAEDLNSGPIKSLILRERQOLRKTY 116
QY 121 SEWQOQLOELTHTSODIEKLKSQYRALARDASQAKRYQASKDKDRKADKXVRSIMKLF 176
DB 121 SEWQOQLOELTHTSODIEKLKSQYRALARDASQAKRYQASKDKDRKADKXVRSIMKLF 176
QY 177 WKLFANHRYVLGVRAQLHGHQHQLLPGLRLSLQDLHEEMACILKEILQETLSSIVQD 236
DB 177 WKLFANHRYVLGVRAQLHGHQHQLLPGLRLSLQDLHEEMACILKEILQETLSSIVQD 236
QY 237 VQDEVALIHEMAAARIQPEAEYQGLROYGSAPRVPCTPFDESILREGEPLREGEOLNE 296
DB 237 VQDEVALIHEMAAARIQPEAEYQGLROYGSAPRVPCTPFDESILREGEPLREGEOLNE 296
QY 297 QNLELTVESVOHTLTSVTDLAVALATEMVPFRQEMVTOLOQELRNEBENTHPRERVOQLGROYL 356
DB 297 QNLELTVESVOHTLTSVTDLAVALATEMVPFRQEMVTOLOQELRNEBENTHPRERVOQLGROYL 356
QY 301 QNLELTVESVOHTLTSVTDLAVALATEMVPFRQEMVTOLOQELRNEBENTHPRERVOQLGROYL 360
DB 301 QNLELTVESVOHTLTSVTDLAVALATEMVPFRQEMVTOLOQELRNEBENTHPRERVOQLGROYL 360
QY 357 ROVLQOGLQVALCSQAKLQAOQELLOTKLEHLGEGEPVLLIQQDRHSTSSSEQER 416
DB 357 ROVLQOGLQVALCSQAKLQAOQELLOTKLEHLGEGEPVLLIQQDRHSTSSSEQER 416
QY 417 EGGRPTLEILKSHISGIRPKFS----- 440
DB 417 EGGRPTLEILKSHISGIRPKFS----- 440
QY 441 -----NLVRLGEGEPSPILPLIDHLSTQOPLTKK 478
DB 441 -----NLVRLGEGEPSPILPLIDHLSTQOPLTKK 478
QY 479 LVRESQKQEVYLVVMMDGHRHFIIOQLDNLVRLGEGEPSPILPLIDHLSTQOPLTKK 538
DB 479 LVRESQKQEVYLVVMMDGHRHFIIOQLDNLVRLGEGEPSPILPLIDHLSTQOPLTKK 538
QY 467 LTKSGVILHRAVPKDKWLVNHEEDLVLSGQIGRNGEVSGLRLADNTLVAVKSCRETL 526
DB 467 LTKSGVILHRAVPKDKWLVNHEEDLVLSGQIGRNGEVSGLRLADNTLVAVKSCRETL 526
QY 539 ITRKSGIVLTRAVALKDKWLVNHEEDLVLSGQIGRNGEVSGLRLADNTLVAVKSCRETL 598
DB 539 ITRKSGIVLTRAVALKDKWLVNHEEDLVLSGQIGRNGEVSGLRLADNTLVAVKSCRETL 598
QY 527 PBDLAKKPELOEALILKQYSHPNIVRLIGVCTQOPRIYVNEVLVOGGPILFPLRTGEGARLV 586
DB 527 PBDLAKKPELOEALILKQYSHPNIVRLIGVCTQOPRIYVNEVLVOGGPILFPLRTGEGARLV 586

```

Db      599 PPELAKFLQEARILKQYHNPNIVRLIGVCTQKQPIYIVMELVQGGDFLSFLRSEGPRLK 658
QY      587 VKTLQMGDAAGMTEYLESKCIHRDLAARNCLVTEKVKLTISDGMREEDGYTAAS 646
Db      659 MKELIKEMEAAGMEYLESKCIHRDLAARNCLVTEKVKLTISDGMREEDGYTAAT 718
QY      647 GGLRQVPVKMTAPALANTGRYSSESVDWMSFGILLMETFSIGASPYPNLSNOQTRFEVKG 706
Db      719 GGMKQIPVKMTAPALANTGRYSSESVDWMSFGILLMETFSIGASPYPNLSNOQTRFEVKG 778
QY      707 GRLPCBELCPDAVFLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 752
Db      779 VRLPEPQCEPDVYRLMQRCEYDPRRRPFSFGAVHODLIAIRKRHR 824

RESULT 5
TVFVPS
protein-tyrosine kinase (BC 2.7.1.112) fps (clone ts) - Fujinami sarcoma virus
C:Species: Fujinami sarcoma virus
C:date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: A26898
R:Chen, L.H.; Hatada, E.; Wheatley, W.; Lee, W.H.
V:tiology 155, 106-119, 1986
A:title: Single amino acid substitution, from Glu-1025 to Asp, of the fps oncogenic prot
A:Reference number: A26898; MUID:87044080; PMID:2877522
A:Accession: A26898
A:Molecule type: DNA
A:Residues: 1-873 <CH>
A:Cross-references: UNIPROT:P00530; GB:M14930; NID:G209668; PID:AAA42403.1; PID:G209689
C:Comment: This protein is synthesized as a gag-fps polyprotein.
C:Genetic:
A:Gene: fps
A:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; tran
F:511-596/Domain: SH2 homology <SH2>
F:510-872/Domain: protein kinase homology <KIN>
F:618-626/Region: protein kinase ATP-binding motif
F:641/Active site: Lys #status predicted

Query Match      66.3%; Score 2570; DB 1; Length 873;
Best Local Similarity 61.4%; Pred. No. 3.5e-92;
Matches 507; Conservative 106; Mismatches 137; Indels 76; Gaps 3;

QY      1 MGFSSELCSPOGHVLCQMGAEFLRLLEGMRKMAQRVKSDBREYAGLHHM-----SLDPS 56
Db      50 MGFGEPLWCPKSGSELRLRLQDSLRLEMLMKKMSQRAKSDREYAGLHHMSQLKQSG 109
QY      57 GGQSRAPSPSPISQSWAETSTQTEGSLRLRQHAEDLNSGPLSKLTLRERQQLRKY 116
Db      110 LGHLRAITHSSQIGESWVLAQTETLSQTLRHAEELAAAGPLAKSLILIRKQQLRKY 169
QY      117 SEWQOQLQELTTHSODIEKLKSQYRALARDSAQAKRYQASXKQDKDRAKQYVRS 176
Db      170 SEWQOQLSQEYAWTTQOEVEKLAQYRSLVRDSTQAKRYQASXKQDKRAKQYVRS 229
QY      177 WKLFANHRVYLVGVAALQHHNQLLPGILRSLODHEMACTIKETLQEYLISL 236
Db      230 SKLYALHNQYLVAVQAAALHHHHYORALPTLHESLYLSLQOEWLVVLEKELGECITSL 289
QY      237 VQDEVAALHREMAAAARLOPEAYOGFLRQYGSAPDVPCVTFDESLLGESEPLRGE 296
Db      290 VQGDVLAIHQEVAAHVEIMDPATEYSSFYQCHRYDSVPVAVTFPDSLLGEASLPGEL 349
QY      297 QNLVELVESVQHTLTSTDELAVATEKVFRRQEMVTQLOQELFNEEBNTHPRRVOLLG 356
Db      350 QNLVELTESVQHSITSLTEELSLASRAVSSKEQVVELQVELRGEELALSPGERVHLGK 409
QY      357 RQVLEALQGLQVLAISOAQALQOQELQTKLHLGSGRPPVTLLODRHSTSSGEOR 416
Db      410 RQGLREHQOQLQGLVCAQAKLQAKORDMLANKLDELSEEPPLALPQEDRQASCTDQK 469
QY      417 EGGRTEPLTELKSHISGIFRPKFS----- 440

```

```

Db      470 SG--VTALKTIKNHISGIFSPFSLPVPVLPEVQKPLCOQAWYGAIPRSEVOELTKY 527
QY      441 -----NLYRLEGSPFILLIHLHLSLTOQP 466
Db      528 SCDPLVRESQKQOEYVLSVMDGQPRHFIQADNLLRLEDGLPTLPIHLQLQSQRP 587
QY      467 LTKSGVVLHRAVPKDKVLANHEDVLAGEQIRGNFGEVSGRLRADNTLVAKSCRETL 526
Db      588 ITRKSGIVLTRAVLQDKVLANHEDVLGEBRIGRNGFGEVSGRLRADNTLPVAKSGRETL 647
QY      527 PDLAKFLQEARILKQYSHPNIVRLIGVCTQKQPIYIVMELVQGGDFLPLRTSGARLR 586
Db      648 PPELAKFLQEARILKQCHPNIVRLIGVCTQKQPIYIVMELVQGGDFLSFLRSEGPRLK 707
QY      587 VKTLQMGDAAGMTEYLESKCIHRDLAARNCLVTEKVKLTISDGMREEDGYTAAS 646
Db      708 MKKLIRKMDNAAAGMEYLESKCIHRDLAARNCLVTEKVKLTISDGMREEDGYTAAT 767
QY      647 GGLRQVPVKMTAPALANTGRYSSESVDWMSFGILLMETFSIGASPYPNLSNOQTRFEVKG 706
Db      768 GGMKQIPVKMTAPALANTGRYSSESVDWMSFGILLMETFSIGASPYPNLSNOQTRFEVKG 827
QY      707 GRLPCBELCPDAVFLMEQCMAYEPGQRPFSFTIYQELQSIKRRHR 752
Db      828 VRLPEPQCEPDVYRLMQRCEYDPRRRPFSFGAVHODLIAIRKRHR 873

RESULT 6
TVFVPS
protein-tyrosine kinase (BC 2.7.1.112) fps - Fujinami sarcoma virus
C:Species: Fujinami sarcoma virus
A:Note: host Gallus gallus (chicken)
C:date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 09-Jul-2004
C:Accession: A00636
R:Shibuya, M.; Hanatua, H.
V:Cell 30, 787-795, 1982
A:title: Nucleotide sequence of Fujinami sarcoma virus: evolutionary relationship of its
A:Reference number: A00636; MUID:83050964; PMID:6291784
A:Accession: A00636
A:Molecule type: genomic RNA
A:Residues: 1-873 <SH1>
A:Cross-references: UNIPROT:P00530
C:Comment: This protein is synthesized as a gag-fps polyprotein.
C:Genetic:
A:Gene: fps
A:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; poly
F:511-596/Domain: SH2 homology <SH2>
F:510-872/Domain: protein kinase homology <KIN>
F:618-626/Region: protein kinase ATP-binding motif
F:641/Active site: Lys #status predicted

Query Match      66.0%; Score 2561; DB 1; Length 873;
Best Local Similarity 61.3%; Pred. No. 7.8e-92;
Matches 506; Conservative 106; Mismatches 138; Indels 76; Gaps 3;

QY      1 MGFSSELCSPOGHVLCQMGAEFLRLLEGMRKMAQRVKSDBREYAGLHHM-----SLDPS 56
Db      50 MGFGEPLWCPKSGTELRLRLQDSLRLEMLMKKMSQRAKSDREYAGLHHMSQLKQSG 109
QY      57 GGQSRAPSPSPISQSWAETSTQTEGSLRLRQHAEDLNSGPLSKLTLRERQQLRKY 116
Db      110 LGHLRAITHSSQIGESWVLAQTETLSQTLRHAEELAAAGPLAKSLILIRKQQLRKY 169
QY      117 SEWQOQLQELTTHSODIEKLKSQYRALARDSAQAKRYQASXKQDKDRAKQYVRS 176
Db      170 SEWQOQLSQEYAWTTQOEVEKLAQYRSLVRDSTQAKRYQASXKQDKRAKQYVRS 229
QY      177 WKLFANHRVYLVGVAALQHHNQLLPGILRSLODHEMACTIKETLQEYLISL 236
Db      230 SKLYALHNQYLVAVQAAALHHHHYORALPTLHESLYLSLQOEWLVVLEKELGECITSL 289

```

```

Oy 237 VODEVVAHREMAAANAARIOPEAEYQGLRQYGAAPDYPCVTTDESLIEGEBLEBCEL 296
   :::::
Db 290 VOEDLALHOKVAHAIVEMIDIPATEYSSFFVOCHRVDSEPPVATPDESILJEEAEMLIEGBEL 349
   :::::
Oy 297 QUNEIYESVOHTLTSVTDELAVALATEWEMFRROEMWTOLOQELRMBEENHTHREBRVOLLGK 356
   :::::
Db 350 QLANETTESVOHSLTSTIEBELLASRAVSSKEORWELQVELRQBELALSGERVHLIGK 409
   :::::
Oy 357 ROVLQEALQGLQVALCSQAKLOAQOELLQTKLEHLGPCEPPVLLLODRHSTSSSEOR 416
   :::::
Db 410 RQGRBRAOQOQGLVCAQAKLOAQORDMLANKLAEIGSEPPPALPLQEDROGSASTQER 469
   :::::
Oy 417 EGCGTPTLEILKSHISGIFRPFSS----- 440
   :::::
Db 470 SG--VTALKTIKHNISGIFSPRFSLPPEVPLIPEVOXKPLCOQAWYHGAIPRSEVOELLKY 527
   :::::
Oy 441 -----NLVRLBEGGFPSTPLIDHLSTQOP 466
   :::::
Db 528 SGDELVRBESQOKOEYVLSVMDGQPRHFTIOADNLTRLBEDGLPTTLPDLIDHLQSORP 587
   :::::
Oy 467 LTKKSGVILHRAVPDDKVLNLHEDLVLEQOIGRGNFGEVFSGRLRADNTLVAVSSCRETL 526
   :::::
Db 588 ITRSGVILTRAVALKDKVNLHEDVLGERIGRGNFGEVFSGRLRADNTPAVAVSCRETL 647
   :::::
Oy 527 PPDLKAKFLQEARILIKOYSHPNIVRLIGVCTOKOPIYIVMELVQGGDFLTLRTEGARLR 586
   :::::
Db 648 PPELKAKFLQEARILIKOCNHPNIVRLIGVCTOKOPIYIVMELVQGGDFLSTRKSGRLK 707
   :::::
Oy 587 VKTLLQWGDAAAGMEYLESKCIHRDLAANCIVTEKNVLIKIDFGSREADAGVVAAS 646
   :::::
Db 708 MKKLIKMMENNAAGMEYLESKCIHRDLAANNCIVTEKNVLIKIDFGSROEBDGVAST 767
   :::::
Oy 647 GGLROVPKWTAPALNATNGRYSSBESDVWSPGILLMETFSLGASBPMLNSQOTREFYEKG 706
   :::::
Db 768 GGMQOIPKWTAPALNATNGRYSSBESDVWSPGILLMEARSLGAVPYANLSNQOTREALEOG 827
   :::::
Oy 707 GRLEPPELCDAVFRLEOCWAYPEGQRPSPSTTYOELQSTIRKGR 752
   :::::
Db 828 VRLPPEQCPEDVYRLMQRCEYDPHRPSPGAVGADIDILIRKGR 873
   :::::

```

```

RESULT 7
TVMVC
protein-tyrosine kinase (EC 2.7.1.112) fes - feline sarcoma virus (strain Gardner-Arnst
C:Species: feline sarcoma virus
C>Note: host Felis sp. (cat)
C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 09-Jul-2004
C:Accession: A00651
R:Hampe, A.; Laprevotte, I.; Gallibert, F.; Fedele, L.A.; Sherr, C.J.
Cell 30, 775-785, 1982
A:Title: Nucleotide sequences of feline retroviral oncogenes (v-fes) provide evidence fo
A:Reference number: A00651; MUID:83050963; PMID:6183005
A:Accession: A00651
A:Molecule type: DNA
A:Residues: 1-609 <HAM>
A:Cross-references: UNIPROT:P00542
C:Comment: This protein is synthesized as a gag-fes polypotein.
C:Gene(s):
C:Gene(s): fes
C:Superfamily: protein-tyrosine kinase fes; protein kinase homology; SH2 homology
C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; trar
F:247-332/Domain: SH2 homology <SH2>
F:346-608/Domain: protein kinase homology <KIN>
F:354-362/Region: protein kinase ATP-binding motif
F:377/Active site: Lys #status predicted

Query Match      63.9%  Score 2479; DB 1; Length 609;
      Bseq Local Similarity  61.7%  Pred. No.7.8e-89;
Matches  507;  Conservative  11;  Mismatches  14;  Indels  290;  Gaps  3;

Oy      1  MGFSSELCSPOGHGVLQOMQAEALRLLEGMRKMAQRVYSAGLLHNSLDSGGQS 60
Db      8  MGFSSELCSPOGHGVEQOMQAEALRLLEGMRKMAQRVYSREYAGLLHNSLDSGG-- 65

```

```

QY 61 RAISPDSPISOSMAITTSQTBEGLSRLLRHOADLNSGPJSKSLLREROQLAKTTSSEOW 120
Db 66 RGTGYSPIOSMAITTSQTBEGLSRLLRHOADLNSGPJSKSLLREROQLAKTTSSEOW 125
QY 121 OOLQOELTKTHSODIEKXKSOYRALARDSAOKKRYKQASKODRKAOKYRSLMKLF 180
Db 126 OOLQOELTKTHNODIEKXKSOYRALARDSAOKKRYKQASKODRKA----- 173
QY 181 AHNRYVLGVRAAQHLHQHHQHLPLGLLRSLDLHEENACTIKELIOEYLAISLVQDS 240
Db 174 ----- 173
QY 241 VVAIHREMAAARARIQPEAEYQGLRQYGSADVPCCVTFDESLEBGEPLBEGELQNB 300
Db 174 ----- 173
QY 301 LVESVQHTLTSVTDELAVATEMVFRRQEMVTLQOQELNBEENTHPRERVOGLGKQVL 360
Db 174 ----- 173
QY 361 QEALQGLVALCSQAKLQAOQELQTKLEHLPGBBPPVLLLODDHSTSSSBOEREGR 420
Db 174 -----KLEQLGGEPPVLLLODDHSTSSSBOEREGR 207
QY 421 TPTLEILKSHISGIRPKFS----- 440
Db 208 TPTLEILKSHISGIRPKFSLPPLQLVPEVOKPLHEOLMYHGALEPRAVEABLLTHSGDF 267
QY 441 -----NLRYLEGEGRPSIPLIDHLSTQOPLTKK 470
Db 268 LVRESQGOEYVLSQLMDGQPRHFIIOASADNLRYPRGGDPFASIPLVLDHLBRQOPLTKK 327
QY 471 SGVLAHRAVPKDKWVLNHEDLVLBQIGRNGFGEVFSGRLRADNTLVAVKSCEETLPPDL 530
Db 328 SGVLAHRAVPKDKWVLNHEDLVLBQIGRNGFGEVFSGRLRADNTLVAVKSCEETLPPDI 387
QY 531 KAKFIOEARILKQYSHPNIVRLIGVCTOKOPRIYIWMELVQGGDFLFLRTEGARLRPKTL 590
Db 388 KAKFIOEARILKQYSHPNIVRLIGVCTOKOPRIYIWMELVQGGDFLFLRTEGARLRPKTL 447
QY 591 LQWGDAAAGMEYLESKCCIRHDLAARNCLEYEKVNLKISDFPMSREEDAGVAAASGLR 650
Db 448 LQWGDAAAGMEYLESKCCIRHDLAARNCLEYEKVNLKISDFPMSREEDAGVAAASGLR 507
QY 651 QVPVKWTAPEALNRYRYSSESQVMSFGIILMETFSIGASPYPNLSNQOTREFYEKGRLP 710
Db 508 QVPVKWTAPEALNRYRYSSESQVMSFGIILMETFSIGASPYPNLSNQOTREFYEKGRLP 567
QY 711 CPFLCPDAVFRLMEQCMAYEPQRPFSFIIYQELQIRKRRH 752
Db 568 CPFLCPDAVFRLMEQCMAYEPQRPFSFIIYQELQIRKRRH 609

```

RESULT 8
 TVMCS
 protein-tyrosine kinase (EC 2.7.1.112) fes - feline sarcoma virus (strain Snyder-Theilen)
 C|Species: feline sarcoma virus
 A|Note: host Felis sp. (cat)
 C|Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #ext_change 09-Jul-2004
 C|Accession: A00652
 R|Hampé, A.; Laprevotte, I.; Galibert, F.; Fedele, L.A.; Sherr, C.J.
 Cell 30, 773-781, 1982
 A|Title: Nucleotide sequences of feline retroviral oncogenes (v-fes) provide evidence for
 A|Reference number: A00651; MUID:83050963; PMID:6183005
 A|Accession: A00652
 A|Molecule type: DNA
 A|Residues: 1-477 <HAM>
 A|Cross-references: UNIPROT:P00543
 C|Comment: This protein is synthesized as a gag-fes polyprotein.
 C|Genetics:
 A|Gene: fes
 C|Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology

C:Keyword: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; tran
 F:115-200/Domain: SH2 homology <SH2>
 F:214-476/Domain: protein kinase homology <KIN>
 F:222-230/Region: protein kinase ATP-binding motif
 F:245/Active site: Lys #status predicted

Query Match 51.0%; Score 1980; DB 1; Length 477;

Best Local Similarity 80.9%; Pred. No. 9, 7e-70;
 Matches 366; Conservative 11; Mismatches 10; Indels 70; Gaps 1;

```

QY 346 HPERVQLGKQVLOEALQGLVALCSQAQKQAOELLQTKLEHLGEPPEPPVLLQDD 405
DB 1 HREGVQLAKKQVLOEALQGLVALCSQAQKQAOELLQTKLEHLGEPPEPPVLLQDD 60
QY 406 RHSTSSSEEREGRGRTPTLEILKSHISGIFRPKES-----NLVRLGEGFPSPIL 440
DB 61 RHSTSSSEEREGRGRTPTLEILKSHISGIFRPKESLPPPLQLVPEVQKPLHQLMYHAL 120
QY 441 -----NLVRLGEGFPSPIL 455
DB 121 PRAEVAELLTHSGDFLVRESQKQEVVLSVMDGQPRHFIQSADLVYRPEGDFASIFL 180
QY 456 LIDHLISTQOPLTKSGVVLHRAVPKQVNLNHEDLVLGEQIGRGNGFVSGRLADNT 515
DB 181 LVDFHLRSQOPLTKSGVILNRAVPKQVNLNHEDLVLGEQIGRGNGFVSGRLADNT 240
QY 516 LVAVVSCTEPLPDLKAKFLQEARILKQVSHPIVLLIGVCTOKOPIYIVMELVQGGDL 575
DB 241 LVAVVSCTEPLPDLKAKFLQEARILKQVSHPIVLLIGVCTOKOPIYIVMELVQGGDL 300
QY 576 TELRTGARLRYKTLQVMDAAAGMEYLESKCCIRHDLAARNCVTEKQVLIKISDFGMS 635
DB 301 TELRTGARLRYKTLQVMDAAAGMEYLESKCCIRHDLAARNCVTEKQVLIKISDFGMS 360
QY 636 REBADGVYAAGGLQVLPYKMTAPEALNTGRYSSESDDVMSFGILLMETFSGLASPYNLS 695
DB 361 REBADGVYAAGGLQVLPYKMTAPEALNTGRYSSESDDVMSFGILLMETFSGLASPYNLS 420
QY 696 NOOTREFVKGGRLPCEPLCPDAVPRLMQCAVYEGORPSFTIYOELISIRKRR 752
DB 421 NOOTREFVKGGRLPCEPLCPDAVPRLMQCAVYEGORPSFTIYOELISIRKRR 477

```

RESULT 9

TYRFP

protein-tyrosine kinase (BC 2.7.1.112) fer - human

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004

C:Accession: A31943; 157573

R:Haoy, Q. L.; Heisterkamp, N.; Groffen, J.

Mol. Cell. Biol. 9, 1587-1593, 1989

A:Title: Isolation and sequence analysis of a novel human tyrosine kinase gene.

A:Reference number: A31943; MUID:89261786; PMID:2725517

A:Accession: A31943

A:Molecule type: mRNA

A:Residues: 1-822 <HAO>

A:Cross-references: UNIPROT:P16591; GB:J03358; NID:9339714; PIDN:AAA61190.1; PID:9339715

C:Genetics: GDB:FER

A:Gene: GDB:FER

A:Cross-references: GDB:125243; OMIM:176942

A:Map position: 5q12-5q14

C:Function:

A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP

C:Superfamily: protein-tyrosine kinase fgs; protein kinase homology; SH2 homology

C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho

F:2-822/Product: protein-tyrosine kinase for #status predicted <MUT>

F:460-546/Domain: SH2 homology <SH2>

F:561-821/Domain: protein kinase homology <KIN>

F:569-577/Region: protein kinase ATP-binding motif

F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

F:591/Active site: Lys #status predicted

Query Match 47.8%; Score 1853; DB 1; Length 822;

Best Local Similarity 45.5%; Pred. No. 1, 3e-64;
 Matches 376; Conservative 150; Mismatches 217; Indels 84; Gaps 9;

```

QY 1 MGFSEELCSPOGHGVLQMOEAEHLRLGMRKMAQRYVSDPEVAGLLHMSLDQSGGS 60
DB 1 MGFSEELCSPOGHGVLQMOEAEHLRLGMRKMAQRYVSDPEVAGLLHMSLDQSGGS 58
QY 1 MGFSEELCSPOGHGVLQMOEAEHLRLGMRKMAQRYVSDPEVAGLLHMSLDQSGGS 58
DB 1 MGFSEELCSPOGHGVLQMOEAEHLRLGMRKMAQRYVSDPEVAGLLHMSLDQSGGS 58
QY 61 RAISPDSPISQMAITSGTEGLSRRLROHAEADLNSGFLSKSLRLREHQRLKTYSEQ 120
DB 59 VQMTAVSNVSKSLMIMQTEQLSRIMKTHADLNSGFLSKSLRLREHQRLKTYSEQ 118
QY 121 QQLQOELTKTHSQDIKTLKSOYRALARDSAOKRYOEA-SKDKRDKAKDRTYVSLMWL 179
DB 119 QQLQOELTKTHSQDIKTLKSOYRALARDSAOKRYOEA-SKDKRDKAKDRTYVSLMWL 178
QY 180 FAHNRRTYLCVRAAOLHONHQLPLGLRLDQLHEMACTLKEILOYEISLVQD 239
DB 179 FAHNRRTYLCVRAAOLHONHQLPLGLRLDQLHEMACTLKEILOYEISLVQD 238
QY 240 EVVAIHREMAAARIOPEAEYQGFRLQYGSAPDVPCVTPDESLEEGEPLEPELQLN 299
DB 239 EVVAIHREMAAARIOPEAEYQGFRLQYGSAPDVPCVTPDESLEEGEPLEPELQLN 298
QY 300 ELTVESVOHTLTSVDELAVALATEVFRQENVTOLOEALNEENTHPR-ERVQLLGRQ 358
DB 299 ELTVESVOHTLTSVDELAVALATEVFRQENVTOLOEALNEENTHPR-ERVQLLGRQ 358
QY 359 VLOEALQGLVALCSQAQKQAOELLQTKLEHLGEPPEPPVLLQDDRHSTSSSEEREG 418
DB 359 VLOEALQGLVALCSQAQKQAOELLQTKLEHLGEPPEPPVLLQDDRHSTSSSEEREG 416
QY 419 GRTPLEILKSHISGIFR-PK-----ESNLYRLGEGFPSPILIDHLSTQ 438
DB 417 GRTPLEILKSHISGIFR-PK-----ESNLYRLGEGFPSPILIDHLSTQ 438
QY 439 -----ESNLYRLGEGFPSPILIDHLSTQ 464
DB 439 -----ESNLYRLGEGFPSPILIDHLSTQ 464
QY 476 KQGDPLVRESHGKPGSYVLVSYSDGQRHFIQVYDMMYRFGTGFSPNIPQIDHYYTK 535
DB 476 KQGDPLVRESHGKPGSYVLVSYSDGQRHFIQVYDMMYRFGTGFSPNIPQIDHYYTK 535
QY 485 OPLTKSGVVLHRAVPKQVNLNHEDLVLGEQIGRGNGFVSGRLADNTLVAVKSR 523
DB 485 OPLTKSGVVLHRAVPKQVNLNHEDLVLGEQIGRGNGFVSGRLADNTLVAVKSR 523
QY 536 QVITTKSGVVLNPLPKQKWLISHEDVILGLLAKNGNVEYVKTGK-DKTSVAVKTK 594
DB 536 QVITTKSGVVLNPLPKQKWLISHEDVILGLLAKNGNVEYVKTGK-DKTSVAVKTK 594
QY 524 ETLPPDLKAKFLQEARILKQVSHPIVLLIGVCTOKOPIYIVMELVQGGDLPTLRTGA 583
DB 524 ETLPPDLKAKFLQEARILKQVSHPIVLLIGVCTOKOPIYIVMELVQGGDLPTLRTGA 583
QY 595 EDLPBELKIKFLQEARILKQVSHPIVLLIGVCTOKOPIYIVMELVQGGDLPTLRTGA 654
DB 595 EDLPBELKIKFLQEARILKQVSHPIVLLIGVCTOKOPIYIVMELVQGGDLPTLRTGA 654
QY 644 AASGGLQVLPYKMTAPEALNTGRYSSESDDVMSFGILLMETFSGLASPYNLSNOQTRFV 703
DB 644 AASGGLQVLPYKMTAPEALNTGRYSSESDDVMSFGILLMETFSGLASPYNLSNOQTRFV 703
QY 715 SSS-GLKQIPKMTAPEALNTGRYSSESDDVMSFGILLMETFSGLASPYNLSNOQTRFV 773
DB 715 SSS-GLKQIPKMTAPEALNTGRYSSESDDVMSFGILLMETFSGLASPYNLSNOQTRFV 773
QY 704 EKGGRLPCEPLCPDAVPRLMQCAVYEGORPSFTIYOELISIRK 750
DB 704 EKGGRLPCEPLCPDAVPRLMQCAVYEGORPSFTIYOELISIRK 750
QY 774 EKGGRLPCEPLCPDAVPRLMQCAVYEGORPSFTIYOELISIRK 820
DB 774 EKGGRLPCEPLCPDAVPRLMQCAVYEGORPSFTIYOELISIRK 820

```

RESULT 10

TYRFP

protein-tyrosine kinase (BC 2.7.1.112) fgs - avian sarcoma virus PRCII

C:Species: avian sarcoma virus PRCII

A:Note: host Gallus gallus (chicken)

C>Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 09-Jul-2004

C:Accession: A00650

R:Huang, C.C.; Hammond, C.; Bishop, J.M.

J. Virol. 50, 125-131, 1984

A:Title: Nucleotide sequence of v-fps in the PRCII strain of avian sarcoma virus.

A:Reference number: A00650; MUID:84138603; PMID:6321783

A:Accession: A00650

A:Molecule type: DNA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

A/Accession: S04328

A/Accession: S04328

A/Accession: S04328

A:Molecule type: mRNA
A:Residues: 1-323 <LET>
A:Cross-references: UNIPROT:P09760; EMBL:X13412; NID:G56169; PIDN:CAA31778.1; PID:G56170
C:Genetics:
A:Gene: Elk
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; tyrosine-speci
F:62-322/Domain: protein kinase homology <KIN>
F:70-78/Region: protein kinase ATP-binding motif

Query Match 28.3%; Score 1099.5; DB 2; Length 323;
Best Local Similarity 66.2%; Pred. No. 8.9e-36;
Matches 206; Conservative 45; Mismatches 57; Indels 3; Gaps 3;

QY 441 NLYRLEGSGPSPSLIDHLSTQOPLTKSGVLLHRAVPKD-KWLVNHEDLVAGQIGR 499
DB 13 NLYRLEGSGSPSLIDHLSTQOPLTKSGVLLHRAVPKD-KWLVNHEDLVAGQIGR 72
QY 500 GNFGVYSGRLRADNTLVAVKSGRETLPPDLKAKFLQEARILKQYSHPNIVRLIGVCTQK 559
DB 73 GNFGVYSGRLK-DKTPVAVKTKCKEDLPQELKIFLQEARILKQYHNPVILKIGVCTQK 131
QY 560 QPIYIVMEIVQGGDELTFTLRTGEGARLVKTLQMGVDAAGMEYLSKCCIHRLAARNC 619
DB 132 QPIYIVMEIVQGGDELTFTLRTGEGARLVKTLQMGVDAAGMEYLSKCCIHRLAARNC 191
QY 620 LVTEKKVLTISDRCGMEADGVYAAAGGLRQVPVKTAPALNNGRYSSESDVMSFGIL 679
DB 192 LVGENNTLTISDRCGMEADGVYAAAGGLRQVPVKTAPALNNGRYSSESDVMSFGIL 250
QY 680 LMETFSLGASPYPNLSNOQTRFVEKGRPLCEPDVFRMLEQCMAYEPQGRPSFT 739
DB 251 LMETFSLGVCYPMGMNQARQEVGRYMSAPQNCPEBFTIMMKCMYKPENRKFSFD 310
QY 740 IYQELQSIKR 750
DB 311 LHKELTAIKKK 321

RESULT 13
149663
tyrosine kinase (fert) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Aug-2004
C:Accession: I49663
R:Fishman, K.; Edman, J.C.; Shackelford, G.M.; Turner, J.A.; Rutter, W.J.; Nir, U.
Mol. Cell. Biol. 10, 146-153, 1990
A:Title: A murine fer testis-specific transcript (fert) encodes a truncated fer protein.
A:Reference number: I49663; MUID:90097822; PMID:2294399
A:Accession: I49663
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-453 <RES>
A:Cross-references: UNIPROT:Q61561; GB:M32054; NID:G193276; PIDN:AAA37617.1; PID:G309235
C:Superfamily: protein kinase homology; SH2 homology
C:Keywords: ATP
F:91-177/Domain: SH2 homology <SH2>
F:192-452/Domain: protein kinase homology <KIN>
F:200-208/Region: protein kinase ATP-binding motif

Query Match 28.3%; Score 1097; DB 2; Length 453;
Best Local Similarity 53.3%; Pred. No. 8.6e-36;
Matches 225; Conservative 63; Mismatches 100; Indels 34; Gaps 8;

QY 348 RERVQLLKKQVLQGLQVALCSQAKLQAOQELL--QTKLEHLGPEPPVLLLODD 405
DB 43 KERKELSLKFSIRHSIAGI-----IKSPKSVLSSTVCDVIVSGRP---LAERD 90
QY 406 RHSTSSSEQ-----REGRRPTLEILKSHISGIR---KFSVLVYLLGE 448
DB 91 WYGAIPRIEAGQLKQGGFLVRESHKGEVYLVSVSGQRHPIFQVNDLYRFEET 150
QY 449 GPFSLILLIDHLSTQOPLTKSGVLLHRAVPKD-KWLVNHEDLVAGQIGRNGPEVPS 507

DB 151 GFSNIPLQIDHFNFKQYITKSGVLLNPIPKDKWLVNHEDLVAGELGKNGEYVK 210
QY 508 GRLRADNTLVAVKSGRETLPPDLKAKFLQEARILKQYSHPNIVRLIGVCTQKPIYIWE 567
DB 211 GLTK-DKTPVAVKTKCKEDLPQELKIFLQEARILKQYHNPVILKIGVCTQKQPIYIWE 269
QY 568 LVQGGDELTFTLRTGEGARLVKTLQMGVDAAGMEYLSKCCIHRLAARNCVTEKNVL 627
DB 270 LVPGGDELTFTLRTGEGARLVKTLQMGVDAAGMEYLSKCCIHRLAARNCVTEKNVL 329
QY 628 KISDFGMSREADGVYAAAGGLRQVPVKTAPALNNGRYSSESDVMSFGILLMETFSIG 687
DB 330 KISDFGMSRQEDGVYSSS-GLKQIPIKWTAEBALNNGRYSSESDVMSFGILLMETFSIG 388
QY 688 ASPYPNLSNOQTRFVEKGRPLCEPDVFRMLEQCMAYEPQGRPSFTIYQELQSI 747
DB 389 VCPYPMGMNQARQEVGRYMSAPQNCPEBFTIMMKCMYKPENRKFSNDLHRELTVI 448
QY 748 RK 749
DB 449 KK 450

RESULT 14
150406
proto-fps protein - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: I50406
R:Peaff, S.L.; Zhou, R.
Virology 146, 307-314, 1985
A:Title: Defining the borders of the chicken proto-fps gene, a precursor of Fujinami sarc
A:Reference number: I50405; MUID:86020620; PMID:2936222
A:Accession: I50406
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-181 <PPA>
A:Cross-references: UNIPROT:Q90943; GB:M11611; NID:G212542; PIDN:AAA9008.1; PID:G212545
C:Genetics:
A:Insertions: 41/2, 94/1, 135/1
C:Superfamily: protein-tyrosine kinase fps; protein kinase homology; SH2 homology
C:Keywords: ATP
F:1-180/Domain: protein kinase homology (fragment) <KIN>

Query Match 18.9%; Score 734; DB 2; Length 181;
Best Local Similarity 73.5%; Pred. No. 3.2e-22;
Matches 133; Conservative 26; Mismatches 22; Indels 0; Gaps 0;

QY 572 GDFLFLRTGRLRVKTLQMGVDAAGMEYLSKCCIHRLAARNCVTEKNVLKISD 631
DB 1 GDFLFLHSBEGHLMKELIKMENAAAGMEYLSKCHIRDLAARNCVTEKNVLKISD 60
QY 632 FQMSREADGVYAAAGGLRQVPVKTAPALNNGRYSSESDVMSFGILLMETFSIGASPY 691
DB 61 FQMSRQEDGVYASPGMKQIPVKTAPALNNGRYSSESDVMSFGILLMEFISGAVPY 120
QY 692 PULSNOQTRFVEKGRPLCEPDVFRMLEQCMAYEPQGRPSFTIYQELQSRKH 751
DB 121 ANLSNOQTRFVEKGRPLCEPDVFRMLEQCMAYEPQGRPSFTIYQELQSRKH 180
QY 752 R 752
DB 181 R 181

RESULT 15
A26132
gab-ab1-pol polyprotein - feline sarcoma virus (strain Hardy-Zuckerman 2) (fragment)
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: feline sarcoma virus
C:Date: 31-Dec-1988 #sequence_revision 09-Sep-1994 #text_change 24-Sep-1999
C:Accession: A26132

R. Bergold, P. J.; Blumenthal, J. A.; D'Andrea, E.; Snyder, H. W.; Lederman, L.; Silverstone, R. J. *Journal of Polymer Science: Part A: Polymer Chemistry* 1997, 35, 1103-1203.

U. WILCOX, 01, 1193-1202, 1987

A/Reference number: A26132; MUID:87141338; PMID:3029415

A;reference number: A26132; MULD:87141338; PMLD:3029415

A;Accession: A26132
A;Molecule type: DNA

A;ResIdues: 1-697 <BER>

A:Cross-references: GB:

C:\Genet1CA: 1 3.....

Dr. Gopal Chandra

А;Вене: гуау-ав1-пу1

Keywords: ATP; oncogene; phosphotransferase; polypeptide; transforming protein; tyrosine kinase; unassigned ser/thr or tyr-specific protein kinases; protein kinase homologs

F;69-117/Domain: SH3 homology <SH3>

F;128-218/Domain: SH2 homology <SH2>

EF:241-501/Domain: protein kinase hol

Fi249-257/Region: protein kinase ATP-binding mo

2-12-77 250/1003-0001: 62000000 1112 21104113
E.272/1004-0001: 1100 40000000 22001000

FILE: LYB HBARCUS predicted

Query match 16.3%; Score 632; DB 2; Length 697;

Best Local Similarity 32.9%; Pred. No. 1e-17;
Matches 163; Conservative 84; Mismatches 147; Indels 102; Gaps 18

348 RERQVLGKRVLQDALQVALCSQAKLQAQQLQTKLHL--GPGPPPVLLQ-- 403

```

Db      17 REHLRY--RQLL---LAGRGARPTNLAQVKQVGGKEENTLAGSENDPLFVALY 71

```

```

404 -----DD-----RHSTSSSEOREGGR-----TP-TLE----- 425

```

Db 72 DEVASGDNTI,STYKGEKI,RYI,GYNHNGEAC,ACTKNAGCWP,SNVYTPBNST,EKHSWYHG 131

[illegible]

```

420 -----LUNSHISGIFKFS-----NLIRLGE 448
      : | | : | : |

```

Db 132 PVSNAAEYLLSSGINGSFLVRESESSIGQRSISLRVEGRVYHYRINTASDGLVSPES 191

```

Qy 449 GPPSIPLLIDH-----LST-QQPLTKKSGVLRHVPK-DKVLNHEDLVGEQIGR 499
    ..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|..|

```

Db 192 RFTLAEIVHHSTVADGLITTLHYAPARKNKPTVGVSPNYDKWEMERIDITMKHKLGG 251

500 GNFGGVFSGRADNTLVAVKSCRE-TLPDLKAKFLQEARILKQYSHPNIVRLIGVCTQ 558

Db 252 GQYEVYEGVWKYSLTVAVKTLKEDTMEVE--EFLKEAAMKEIKHPNLVOLLGVCIR 308

00 559 K0P1Y1M6I.V0G0D6I.T6I.P7E6C0R1.BYK--T11.0M/CDA0006NEV7.ESK0C1H0D1.A0 616

[illegible]

DB 309 EPPF111EFMI5NLEDDYLR-ECNRQZVNAVLELYMAIQLISSAMEYLEKKNFIHRDLFA 36/

617 RNC LVTEK N V L K I S D F G M S R E A D C V Y A S G G L R Q V P V K T I A P E A L N Y G R Y S S E S D W S F 676

Db 368 RNCVGENHLVKVADGSLMTGDTYYAHAGTK-FLIKMTAPESLAYNKFISIKSDVWAF 426

677 GILMETFSLGASPYPNLSNQOTREFVEKGRLPCPELCPDAVRLMEQCWAYEPGQPS 736

Db 427 GVLLMEIATYGMSPYPCGIDLSQVYELLEKDYRMERPECGPEEKVYELMRACMQWNPSPDRPA 486

Search completed: March 18, 2005, 16:00:10
Job time : 50 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 15:41:45 ; Search time 170 Seconds
(without alignments)
1710.846 Million cell updates/sec

Title: US-10-660-763-2

Perfect score: 3879

Sequence: 1 MGFSEELCSPOGHGVLQGMQ.....QRPSTIYQELQSRKRR 752

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp19808:***
2: geneseqp19908:***
3: geneseqp20008:***
4: geneseqp20018:***
5: geneseqp20028:***
6: geneseqp20038:***
7: geneseqp20038:***
8: geneseqp20048:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3879	100.0	752	5	AB07354 Human pro
2	3879	100.0	752	6	AB07354 Human pro
3	3879	100.0	752	8	AB07354 Human pro
4	3879	100.0	752	8	AB07354 Human pro
5	3879	100.0	752	8	AB07354 Human pro
6	3879	100.0	752	8	AB07354 Human pro
7	3879	100.0	752	8	AB07354 Human pro
8	3879	100.0	752	8	AB07354 Human pro
9	3879	100.0	752	8	AB07354 Human pro
10	3879	100.0	752	8	AB07354 Human pro
11	3879	100.0	752	8	AB07354 Human pro
12	3879	100.0	752	8	AB07354 Human pro
13	3879	100.0	752	8	AB07354 Human pro
14	3879	100.0	752	8	AB07354 Human pro
15	3879	100.0	752	8	AB07354 Human pro
16	3879	100.0	752	8	AB07354 Human pro
17	3879	100.0	752	8	AB07354 Human pro
18	3879	100.0	752	8	AB07354 Human pro
19	3879	100.0	752	8	AB07354 Human pro
20	3879	100.0	752	8	AB07354 Human pro
21	3879	100.0	752	8	AB07354 Human pro
22	3879	100.0	752	8	AB07354 Human pro
23	3879	100.0	752	8	AB07354 Human pro
24	3879	100.0	752	8	AB07354 Human pro
25	3879	100.0	752	8	AB07354 Human pro

26	642	16.6	256	4	AA009686	AA009686 Human pol
27	642	16.6	280	8	AD114134	AD114134 Novel hum
28	642	16.6	280	8	AD114134	AD114134 Novel hum
29	630.5	16.3	1148	5	ABP65122	ABP65122 Hypoxia-r
30	628	16.2	1130	6	ABG76258	ABG76258 Human Bcr
31	626	16.1	1130	6	ABU10169	ABU10169 Human mut
32	626	16.1	1130	6	ABG76236	ABG76236 Human Bcr
33	626	16.1	1130	6	ABG76288	ABG76288 Human Bcr
34	626	16.1	1130	6	ABG76248	ABG76248 Human Bcr
35	626	16.1	1130	6	ABG76264	ABG76264 Human Bcr
36	625	16.1	1130	6	ABU10166	ABU10166 Human mut
37	625	16.1	1130	6	ABG76222	ABG76222 Human Bcr
38	625	16.1	1130	6	ABG76259	ABG76259 Human Bcr
39	625	16.1	1130	6	ABG76214	ABG76214 Human Bcr
40	625	16.1	1130	6	ABG76210	ABG76210 Human Bcr
41	624.5	16.1	1130	6	AAE34489	AAE34489 Human Abl
42	624.5	16.1	1148	7	ADJ68940	ADJ68940 Human hea
43	624	16.1	1130	6	ABU10164	ABU10164 Human mut
44	624	16.1	1130	6	ABU10182	ABU10182 Human mut
45	624	16.1	1130	6	ABU10168	ABU10168 Human mut

ALIGNMENTS

RESULT 1
ID ABB07354 standard; protein; 752 AA.
XX
AC ABB07354;
XX
DT 09-APR-2002 (first entry)
XX
DE Human proto-oncogene tyrosine kinase.
XX
KW Proto-oncogene tyrosine kinase; poTK; tumour; cytostatic; anti-leukemic;
KW gene therapy; protein therapy; vaccine; enzymatic-inhibition; human;
KW anti-kinase.
XX
OS Homo sapiens.
XX
PN US6340584-B1.
XX
PD 22-JUN-2002.
XX
PF 27-MAR-2001; 2001US-00817180.
XX
PR 27-MAR-2001; 2001US-00817180.
XX
PA (PEKE) PE CORP NY.
XX
PI Gan W, Ye J, Di Francesco V, Beasley EM;
XX
DR N-PSDB; ABA94500, ABA94501.
XX
PT Nucleic acids encoding a proto-oncogene tyrosine kinase, useful for the
XX prevention, diagnosis and treatment of e.g. leukemia and lung tumors.
XX
PS Claim 1; Fig 2A; 49p; English.
XX
CC The invention provides isolated nucleic acid sequences encoding a proto-
CC oncogene tyrosine kinase (poTK). The poTK polynucleotides and protein may
CC be used in the prevention, diagnosis and treatment of diseases associated
CC with inappropriate poTK expression, such as lung and kidney tumours,
CC leukemia and stomach adenocarcinoma. poTK may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of poTK by expressing
CC inactive proteins or to supplement the patient's own production of poTK.
CC The encoded poTK may be used as an antigen in the production of poTK
CC antibodies against poTK and in assays to identify modulators of poTK
CC expression and activity. The anti-poTK antibodies and antagonists may be
CC used to down regulate expression and activity and as diagnostic agents

QY 241 VVAIHREMAAAARIOPEAEYQFLROYGSADVPVCVTFDSESLIEGEPLEBGELOLNE 300
 DB 241 VVAIHREMAAAARIOPEAEYQFLROYGSADVPVCVTFDSESLIEGEPLEBGELOLNE 300
 QY 301 LTVESVQHTLTSVTDELAVALATEMVFRRQEMVTQLOQELNNEBENTHPRERVOILGKROYL 360
 DB 301 LTVESVQHTLTSVTDELAVALATEMVFRRQEMVTQLOQELNNEBENTHPRERVOILGKROYL 360
 QY 361 QEALQGLQVALCSQAKLQAOQELQTKLHLGEPPEPVLILQDDHSTSSSEQEREGR 420
 DB 361 QEALQGLQVALCSQAKLQAOQELQTKLHLGEPPEPVLILQDDHSTSSSEQEREGR 420
 QY 421 TPTLEILKSHISGIFRPKFSNLYRLEGEPSIPILIDHLSTQOPLTKKSGVHLRAVP 480
 DB 421 TPTLEILKSHISGIFRPKFSNLYRLEGEPSIPILIDHLSTQOPLTKKSGVHLRAVP 480
 QY 481 KDKWVLMHEDLVLEQIGRNGEVEFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR 540
 DB 481 KDKWVLMHEDLVLEQIGRNGEVEFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR 540
 QY 541 LKQYSHPNIVRLIGVCTQKOPITYIMELVQGGDFLTPLETEGARLVKTLQWVGDAAG 600
 DB 541 LKQYSHPNIVRLIGVCTQKOPITYIMELVQGGDFLTPLETEGARLVKTLQWVGDAAG 600
 QY 601 MEYLESKCCIHRLDAAARNCLVTEKNVLTSDFGMSREEDGVYAAAGSLRQVPVKMTAB 660
 DB 601 MEYLESKCCIHRLDAAARNCLVTEKNVLTSDFGMSREEDGVYAAAGSLRQVPVKMTAB 660
 QY 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYPLNSQOTREFEVEKGRLLPCBELCPDAVF 720
 DB 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYPLNSQOTREFEVEKGRLLPCBELCPDAVF 720
 QY 721 RLMEQCMAYEPGQRPFSFTIYOELQSIKRRHR 752
 DB 721 RLMEQCMAYEPGQRPFSFTIYOELQSIKRRHR 752

RESULT 3
 ADM28581
 ID ADM28581 standard; protein; 752 AA.
 XX
 AC ADM28581;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Human proto-oncogene tyrosine kinase.
 XX
 KW gene therapy; human; proto-oncogene; tyrosine kinase; cancer; enzyme.
 XX
 OS Homo sapiens.
 XX
 PN US2004063130-A1.
 XX
 PD 01-APR-2004.
 XX
 PF 12-SEP-2003; 2003US-00660763.
 XX
 PR 27-MAR-2001; 2001US-00817180.
 PR 06-DEC-2001; 2001US-00003295.
 XX
 PA (APPL-) APPLERA CORP.
 XX
 PI Gan W, Ye J, Di Francesco V, Beasley EW;
 XX
 DR MPI; 2004-282461/26.
 DR N-PSDB; ADM28580, ADM28582.
 XX
 PT New human kinase peptides, useful for preparing a composition for
 XX treating a disease or condition mediated by human kinases.
 PS Claim 1; SEQ ID NO 2; 53pp; English.
 CC The invention relates to a new isolated human proto-oncogene tyrosine

CC kinase peptide or its allelic variant, orthologue or fragment. The
 CC peptide is useful for preparing a composition for treating a disease or
 CC condition mediated by a human kinase protein e.g. cancer. The present
 CC sequence represents the amino acid sequence of the human proto-oncogene
 CC tyrosine kinase.
 XX
 SQ Sequence 752 AA;
 QY 1 MFSSSELSPQGHVQLQOQAEALRLLEGMRKMAQVVSDEYAGLHHMSIQDSGGOS 60
 DB 1 MFSSSELSPQGHVQLQOQAEALRLLEGMRKMAQVVSDEYAGLHHMSIQDSGGOS 60
 QY 61 RAISPSPISQSWAETTSQTEGLSRLLRQHADLNSGPLSKSLIRERQQLRTYSBOM 120
 DB 61 RAISPSPISQSWAETTSQTEGLSRLLRQHADLNSGPLSKSLIRERQQLRTYSBOM 120
 QY 121 QOLQOELTTHSQDIEKLKQYRALARDSAOKKRYQASKQKDKDAOKYVRSIMKLF 180
 DB 121 QOLQOELTTHSQDIEKLKQYRALARDSAOKKRYQASKQKDKDAOKYVRSIMKLF 180
 QY 181 AHNHNRVVLGVRAQLHGHQHLPLGLRLSLQDLHEBACTLKETLOEYLSISLVODE 240
 DB 181 AHNHNRVVLGVRAQLHGHQHLPLGLRLSLQDLHEBACTLKETLOEYLSISLVODE 240
 QY 241 VVAIHREMAAAARIOPEAEYQFLROYGSADVPVCVTFDSESLIEGEPLEBGELOLNE 300
 DB 241 VVAIHREMAAAARIOPEAEYQFLROYGSADVPVCVTFDSESLIEGEPLEBGELOLNE 300
 QY 301 LTVESVQHTLTSVTDELAVALATEMVFRRQEMVTQLOQELNNEBENTHPRERVOILGKROYL 360
 DB 301 LTVESVQHTLTSVTDELAVALATEMVFRRQEMVTQLOQELNNEBENTHPRERVOILGKROYL 360
 QY 361 QEALQGLQVALCSQAKLQAOQELQTKLHLGEPPEPVLILQDDHSTSSSEQEREGR 420
 DB 361 QEALQGLQVALCSQAKLQAOQELQTKLHLGEPPEPVLILQDDHSTSSSEQEREGR 420
 QY 421 TPTLEILKSHISGIFRPKFSNLYRLEGEPSIPILIDHLSTQOPLTKKSGVHLRAVP 480
 DB 421 TPTLEILKSHISGIFRPKFSNLYRLEGEPSIPILIDHLSTQOPLTKKSGVHLRAVP 480
 QY 481 KDKWVLMHEDLVLEQIGRNGEVEFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR 540
 DB 481 KDKWVLMHEDLVLEQIGRNGEVEFSGRLRADNTLVAVKSCRETLPPDLKAKFLQEAR 540
 QY 541 LKQYSHPNIVRLIGVCTQKOPITYIMELVQGGDFLTPLETEGARLVKTLQWVGDAAG 600
 DB 541 LKQYSHPNIVRLIGVCTQKOPITYIMELVQGGDFLTPLETEGARLVKTLQWVGDAAG 600
 QY 601 MEYLESKCCIHRLDAAARNCLVTEKNVLTSDFGMSREEDGVYAAAGSLRQVPVKMTAB 660
 DB 601 MEYLESKCCIHRLDAAARNCLVTEKNVLTSDFGMSREEDGVYAAAGSLRQVPVKMTAB 660
 QY 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYPLNSQOTREFEVEKGRLLPCBELCPDAVF 720
 DB 661 ALNYGRYSSESDVWSFGILLMETFSLGASPPYPLNSQOTREFEVEKGRLLPCBELCPDAVF 720
 QY 721 RLMEQCMAYEPGQRPFSFTIYOELQSIKRRHR 752
 DB 721 RLMEQCMAYEPGQRPFSFTIYOELQSIKRRHR 752

RESULT 4
 ABB07355
 ID ABB07355 standard; protein; 822 AA.
 XX
 AC ABB07355;
 XX
 DT 09-APR-2002 (first entry)
 XX

DE Human homologue of feline sarcoma virus FES oncogene product.
 XX
 XX Proto-oncogene tyrosine kinase; potk; tumour; cytostatic; anti-leukemic;
 KM gene therapy; protein therapy; vaccine; enzymatic-inhibition; human;
 KM anti-kinase; oncogene; FES.
 XX
 XX Homo sapiens.
 XX
 XX US6340584-B1.
 XX
 XX 22-JAN-2002.
 XX
 XX 27-MAR-2001; 2001US-00817180.
 XX
 XX 27-MAR-2001; 2001US-00817180.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Gan W, Ye J, Di Francesco V, Beasley EM;
 XX
 XX WPI; 2002-138497/18.
 XX
 XX Nucleic acids encoding a proto-oncogene tyrosine kinase, useful for the
 PT prevention, diagnosis and treatment of e.g. leukemia and lung tumors.
 XX
 XX Disclosure; Fig 2C; 49pp; English.
 XX
 XX The invention provides isolated nucleic acid sequences encoding a proto-
 CC oncogene tyrosine kinase (potk). The potk polynucleotides and protein may
 CC be used in the prevention, diagnosis and treatment of diseases associated
 CC with inappropriate potk expression, such as lung and kidney tumors,
 CC leukemia and stomach adenocarcinoma. potk may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of potk by expressing
 CC inactive proteins or to supplement the patient's own production of potk.
 CC The encoded potk may be used as an antigen in the production of
 CC antibodies against potk and in assays to identify modulators of potk
 CC expression and activity. The anti-potk antibodies and antagonists may be
 CC used to down regulate expression and activity and as diagnostic agents
 CC for detecting the presence of potk in samples. The present sequence
 CC represents the human homologue of feline sarcoma virus FES oncogene
 CC product
 CC
 XX
 XX Sequence 822 AA;
 XX
 XX
 XX Query Match 98.8%; Score 3834; DB 5; Length 822;
 XX Best Local Similarity 91.5%; Pred. No. 1.5e-276;
 XX Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;
 XX
 QY 1 MGFSSELCSPOGHGVQOMQOAEELRLLEGMRKMAQORVSDREYAGLHHMSIQDSGQS 60
 DB 1 MGFSSELCSPOGHGVQOMQOAEELRLLEGMRKMAQORVSDREYAGLHHMSIQDSGQS 60
 QY 61 RAISPDSPISQSWAEITSGTEGLSRLLRQHAEDLNSGPI.SKLSLLIREBQOLAKTYSQW 120
 DB 61 RAISPDSPISQSWAEITSGTEGLSRLLRQHAEDLNSGPI.SKLSLLIREBQOLAKTYSQW 120
 QY 121 QOLQOELITKTHSGDITKLSQYALARDNAQAKRKQESKDPDRKADKYRSLWKLP 180
 DB 121 QOLQOELITKTHSGDITKLSQYALARDNAQAKRKQESKDPDRKADKYRSLWKLP 180
 QY 121 QOLQOELITKTHSGDITKLSQYALARDNAQAKRKQESKDPDRKADKYRSLWKLP 180
 DB 121 QOLQOELITKTHSGDITKLSQYALARDNAQAKRKQESKDPDRKADKYRSLWKLP 180
 QY 181 AHHNRVVLGVRQAQLHHQHHLPLGLRLSLQDLHEMACTIKELLOEYLETSSIVQOE 240
 DB 181 AHHNRVVLGVRQAQLHHQHHLPLGLRLSLQDLHEMACTIKELLOEYLETSSIVQOE 240
 QY 241 VVAIHREMAAAARIQPEAEYQGLRQYSGAPDPVPCVTFDESILBEGEPLEBGLQUNE 300
 DB 241 VVAIHREMAAAARIQPEAEYQGLRQYSGAPDPVPCVTFDESILBEGEPLEBGLQUNE 300
 QY 301 LTVESVQHTLTSTTDELAVATENVFRQEWVTQLOQLNEBENTPRERVOLLRGQVYL 360
 DB 301 LTVESVQHTLTSTTDELAVATENVFRQEWVTQLOQLNEBENTPRERVOLLRGQVYL 360

QY 361 QEALQGLQVALCSQAKLQAQOELLQTKLEHLGPGEPPEPYLLIQDDRHSSTSSQEREGGR 420
 DB 361 QEALQGLQVALCSQAKLQAQOELLQTKLEHLGPGEPPEPYLLIQDDRHSSTSSQEREGGR 420
 QY 421 TPTLEILKSHISGIRPFPFS----- 440
 DB 421 TPTLEILKSHISGIRPFPFS----- 440
 QY 441 -----NLRYLEGEGFPISIPLIDHILSTQOPLTK 470
 DB 481 LVRESQGOEYVLSVLMGDLPHNFIIOSLIDNLYRLEGEGFPISIPILIDHILSTQOPLTK 540
 QY 471 SGVVLHRAVVPKDKWVLNHDVLGEQIGKNGFEVFSGRLRADNTLVAVKSCRETLPPDL 530
 DB 541 SGVVLHRAVVPKDKWVLNHDVLGEQIGKNGFEVFSGRLRADNTLVAVKSCRETLPPDL 600
 QY 531 KAKPLQEARILKQYSHPNIVRLIGVCTQKOPYYIYMEVQGGDFLFTLEGARLRVKT 590
 DB 601 KAKPLQEARILKQYSHPNIVRLIGVCTQKOPYYIYMEVQGGDFLFTLEGARLRVKT 660
 QY 591 LQWGDAAAGMEYLESKCCIRHDLAARNCVTEKRVLTKIDRGMREADGVYAASGGR 650
 DB 661 LQWGDAAAGMEYLESKCCIRHDLAARNCVTEKRVLTKIDRGMREADGVYAASGGR 720
 QY 651 QVPVMTAPEALNRYGSSSDVMSFGILLMETFSLGASPPNLSNQOTREFVEKGRLLP 710
 DB 721 QVPVMTAPEALNRYGSSSDVMSFGILLMETFSLGASPPNLSNQOTREFVEKGRLLP 780
 QY 711 CPELCPPDAVFRMECCMAVEPQQRPSFTIYVELSIRKRR 752
 DB 781 CPELCPPDAVFRMECCMAVEPQQRPSFTIYVELSIRKRR 822
 XX
 XX RESULT 5
 XX ABB99047
 XX ID ABB99047 standard; protein; 822 AA.
 XX
 XX ABB99047;
 XX
 XX 24-JAN-2003 (first entry)
 XX
 XX
 XX Human kinase related amino acid sequence.
 XX
 XX
 XX Human; kinase; proto-oncogene tyrosine kinase; antiarteriosclerotic;
 KM cytostatic; antiinflammatory; antiapoptotic; gene therapy;
 KM protein kinase; drug screening assay; tissue typing; chromosome 15;
 KM pharmacogenomic analysis; inflammation; cancer; leukaemia; lung tumours;
 KM kidney tumour; stomach adenocarcinoma; arteriosclerosis; psoriasis.
 XX
 XX Homo sapiens.
 XX
 XX WO200277191-A2.
 XX
 XX 03-OCT-2002.
 XX
 XX 27-MAR-2002; 2002WO-US009325.
 XX
 XX 27-MAR-2001; 2001US-00817180.
 XX
 XX 06-DEC-2001; 2001US-00003295.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Gan W, Ye J, Di Francesco V, Beasley EM;
 XX
 XX WPI; 2003-029927/02.
 XX
 XX New human kinase protein, useful for treating or diagnosing disorders
 PT associated with an absence of, inappropriate, or unwanted expression of
 PT the protein, e.g. inflammation or cancer, in drug screening assays and
 PT pharmacogenomics.
 XX
 XX Disclosure; Page 74-75; 75pp; English.
 XX

CC The invention relates to a newly isolated peptide sequence of a human
 CC kinase that is related to the proto-oncogene tyrosine kinase subfamily.
 CC The activity of the kinase of the invention may be described as,
 CC cytoleptic, antiarteriosclerotic, antiinflammatory and antiproliferative.
 CC peptides of the invention are useful in assays to determine the
 CC biological activity of the protein, in drug screening assays, tissue
 CC typing and pharmacogenomic analysis. They are also useful in treating or
 CC diagnosing disorders characterized by an absence of, inappropriate, or
 CC unwanted expression of the protein, such as inflammation, cancer (e.g.
 CC leukemia, lung tumours, kidney tumours or stomach adenocarcinoma),
 CC arteriosclerosis, and psoriasis. Nucleic acid molecules of the invention
 CC are useful as probes, primers and chemical intermediates in biological
 CC assays. The peptide and nucleic acid sequences are useful as models for
 CC the development of human therapeutic targets, aid in the identification
 CC of therapeutic proteins and serve as targets for the development of human
 CC therapeutic agents that modulate kinase protein activity in cells and
 CC tissues that express the protein. The protein of the invention may also
 CC be useful in gene therapy. The gene encoding the protein of the invention
 CC has been localised to human chromosome 15. The current sequence
 CC represents an amino acid sequence relative to the human kinase sequence
 CC of the invention. NOTE: This sequence is not further mentioned in the
 CC specification
 CC XX
 CC XX

SQ Sequence 822 AA;

Query Match 98.8%; Score 3834; DB 6; Length 822;

Best Local Similarity 91.5%; Pred. No. 1.5e-276;

Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGFSSSLCSPQGHGVLCQOEAEELRLLEGKMKMAQRYKSDREYAGLLHMSLQDSGGQS 60
 DB 1 MGFSSSLCSPQGHGVLCQOEAEELRLLEGKMKMAQRYKSDREYAGLLHMSLQDSGGQS 60
 QY 61 RAISPSPISSQSWAEITSTQTEGSRLLRQAEEDLNSGPKSLILRERQQLKRTYSEOW 120
 DB 61 RAISPSPISSQSWAEITSTQTEGSRLLRQAEEDLNSGPKSLILRERQQLKRTYSEOW 120
 QY 121 QQLQOELTTHSGDIEKLKSOYRALARDSQAQARKYQASKDKDRDKAKDKYRSIMKLF 180
 DB 121 QQLQOELTTHSGDIEKLKSOYRALARDSQAQARKYQASKDKDRDKAKDKYRSIMKLF 180
 QY 122 QQLQOELTTHSGDIEKLKSOYRALARDSQAQARKYQASKDKDRDKAKDKYRSIMKLF 180
 DB 122 QQLQOELTTHSGDIEKLKSOYRALARDSQAQARKYQASKDKDRDKAKDKYRSIMKLF 180
 QY 181 AHNHNRVVLGVRACQLHHQHOLLLPGLRLSLQDLHHEMACILKEILOEYLETSSLYODE 240
 DB 181 AHNHNRVVLGVRACQLHHQHOLLLPGLRLSLQDLHHEMACILKEILOEYLETSSLYODE 240
 QY 241 VVAIHREMAAAARIQPEAEYQGLRQYGSAPVPCVTFDESILBEGEPLEPGELOLNE 300
 DB 241 VVAIHREMAAAARIQPEAEYQGLRQYGSAPVPCVTFDESILBEGEPLEPGELOLNE 300
 QY 301 LTVESVQHTLTSTYDELAVATEVFRROEMWTOLQOELRNEBENTHPRERVOQLGRQVL 360
 DB 301 LTVESVQHTLTSTYDELAVATEVFRROEMWTOLQOELRNEBENTHPRERVOQLGRQVL 360
 QY 361 QEALQGLQVALGSOAKLQAOQELLOTKLHILGEPPEPVLLQDDHSTSSSQBERGGR 420
 DB 361 QEALQGLQVALGSOAKLQAOQELLOTKLHILGEPPEPVLLQDDHSTSSSQBERGGR 420
 QY 421 TPTLEILKSHISGIFRPKFS----- 440
 DB 421 TPTLEILKSHISGIFRPKFS----- 440
 QY 441 -----NLYRLEEGEPSPILLDHLLSTFOQLTKK 470
 DB 481 LVRESQKQKQRYVLMVMDGLPRHFIQSLDNLVRLBEGEPSPILLDHLLSTFOQLTKK 540
 QY 471 SGVVLRAVPKDKWLVNHEDLVYGEQIGRNGFESVSGRLRANTLVAVVSCETLPPDL 560
 DB 541 SGVVLRAVPKDKWLVNHEDLVYGEQIGRNGFESVSGRLRANTLVAVVSCETLPPDL 600
 QY 531 KAKFLEARILKQYSHPNIVRLIGVCTQKQPIYIYVWELVGGDFLFLRTGARLAVKTL 590
 DB 601 KAKFLEARILKQYSHPNIVRLIGVCTQKQPIYIYVWELVGGDFLFLRTGARLAVKTL 660

QY 591 LQWVGDAAGMEYLESKCIHRDLARNCLVTEKYNLXISDFEGMSREBAGVYAAAGGLR 650
 DB 661 LQWVGDAAGMEYLESKCIHRDLARNCLVTEKYNLXISDFEGMSREBAGVYAAAGGLR 720
 QY 651 QVPVKMTAPEALNAGRYSESSESDVSPGILLMETFISGASPPYNI,SNQOTREPVKGGRLP 710
 DB 721 QVPVKMTAPEALNAGRYSESSESDVSPGILLMETFISGASPPYNI,SNQOTREPVKGGRLP 780
 QY 711 CPELCDAVFRIMEQCMAYEPGQRPSEFTTYOELQIRKRR 752
 DB 781 CPELCDAVFRIMEQCMAYEPGQRPSEFTTYOELQIRKRR 822

RESULT 6

ID ADM28583 standard; protein; 822 AA.

AC ADM28583;

DT 17-JUN-2004 (first entry)

DE Human V-FES/FPS protein.

KM Gene therapy; human; proto-oncogene; tyrosine kinase; cancer; enzyme;

KW V-FES/FPS.

OS Homo sapiens.

PN US2004063130-A1.

PD 01-APR-2004.

PF 12-SEP-2003; 2003US-00660763.

PR 27-MAR-2001; 2001US-00817180.

PR 06-DEC-2001; 2001US-00003295.

PA (APPL-) APPLERA CORP.

PI Gan W, Ye J, Di Francesco V, Beasley EM;

DR WPI; 2004-282461/26.

XX New human kinase peptides, useful for preparing a composition for

PT treating a disease or condition mediated by human kinases.

PS Disclosure; SEQ ID NO 4; 53pp; English.

CC The invention relates to a new isolated human proto-oncogene tyrosine

CC kinase peptide or its allelic variant, orthologue or fragment. The

CC peptide is useful for preparing a composition for treating a disease or

CC condition mediated by a human kinase protein e.g. cancer. The present

CC sequence represents the amino acid sequence of the human V-FES/FPS

CC protein.

CC XX

SQ Sequence 822 AA;

Query Match 98.8%; Score 3834; DB 8; Length 822;

Best Local Similarity 91.5%; Pred. No. 1.5e-276;

Matches 752; Conservative 0; Mismatches 0; Indels 70; Gaps 1;

QY 1 MGFSSSLCSPQGHGVLCQOEAEELRLLEGKMKMAQRYKSDREYAGLLHMSLQDSGGQS 60
 DB 1 MGFSSSLCSPQGHGVLCQOEAEELRLLEGKMKMAQRYKSDREYAGLLHMSLQDSGGQS 60
 QY 61 RAISPSPISSQSWAEITSTQTEGSRLLRQAEEDLNSGPKSLILRERQQLKRTYSEOW 120
 DB 61 RAISPSPISSQSWAEITSTQTEGSRLLRQAEEDLNSGPKSLILRERQQLKRTYSEOW 120
 QY 121 QQLQOELTTHSGDIEKLKSOYRALARDSQAQARKYQASKDKDRDKAKDKYRSIMKLF 180
 DB 121 QQLQOELTTHSGDIEKLKSOYRALARDSQAQARKYQASKDKDRDKAKDKYRSIMKLF 180

QY 181 AHNHRYVLGVRAAQLHHQHLLPGLRLSLQDLHEENACTIKETLQETLEISLVDDE 240
 DB 181 AHNHRYVLGVRAAQLHHQHLLPGLRLSLQDLHEENACTIKETLQETLEISLVDDE 240
 QY 241 VVAIHREMAAAARIQPEAEYQGFRLQYGSADVPVPCVTFDESLEEGEPLEGEQLQNE 300
 DB 241 VVAIHREMAAAARIQPEAEYQGFRLQYGSADVPVPCVTFDESLEEGEPLEGEQLQNE 300
 QY 301 LTVESVQHTLTSVTDELAVALATEMVFRRQEMVNTQLOELNNEENHTPRERVOQLGKROYL 360
 DB 301 LTVESVQHTLTSVTDELAVALATEMVFRRQEMVNTQLOELNNEENHTPRERVOQLGKROYL 360
 QY 361 QEALQGLQVALCSQAKLQAOQELQTKLEHLGPGEPPVLLQODDRHSTSSSQEREGGR 420
 DB 361 QEALQGLQVALCSQAKLQAOQELQTKLEHLGPGEPPVLLQODDRHSTSSSQEREGGR 420
 QY 421 TPTLEILKSHISGIFRPKFS----- 440
 DB 421 TPTLEILKSHISGIFRPKFS----- 440
 QY 441 -----NLVRLSEGGPSPILIDHLSTQOPLTKK 470
 DB 481 LVRESQKQBYVLVSLMDGLPRHFIITQSLDNLVRLSEGGPSPILIDHLSTQOPLTKK 540
 QY 471 SGVVLHRAVPKDKMVLNHEDLVLGEQIGRNGFGEVSGRLRADNTLVAVKSCHETLPPDL 530
 DB 541 SGVVLHRAVPKDKMVLNHEDLVLGEQIGRNGFGEVSGRLRADNTLVAVKSCHETLPPDL 600
 QY 531 KAKFLOBARILKQYSHPNVRLIGVCTQKOPITYINVELVOGGPPLFPLRTGEGARLVKTL 590
 DB 601 KAKFLOBARILKQYSHPNVRLIGVCTQKOPITYINVELVOGGPPLFPLRTGEGARLVKTL 660
 QY 591 LQWVGDAAGMEYLESKCCIHRLDAAARNCLVTEKNVLKISDFGMSREBAGVVAASGGRL 650
 DB 661 LQWVGDAAGMEYLESKCCIHRLDAAARNCLVTEKNVLKISDFGMSREBAGVVAASGGRL 720
 QY 651 QVPVKMTABEALNYGRYSSESDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRLLP 710
 DB 721 QVPVKMTABEALNYGRYSSESDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRLLP 780
 QY 711 CPRLCPDAVFRMLEQCMAYEPGQRPSPSTTYQELQSIIRKRR 752
 DB 781 CPRLCPDAVFRMLEQCMAYEPGQRPSPSTTYQELQSIIRKRR 822
 RESULT 7
 AAU32655 standard; protein; 2202 AA.
 AAU32655;
 18-DEC-2001 (first entry)
 Novel human secreted protein #3146.
 Human; vaccination; gene therapy; nutritional supplement;
 stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 Homo sapiens.
 MO200179449-A2.
 25-OCT-2001.
 16-APR-2001; 2001MO-US008656.
 18-APR-2000; 2000US-00552929.
 26-JAN-2001; 2001US-00770160.
 (HYSE-) HYSEQ INC.
 Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-611725/70.
 DR Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX
 XX
 XX
 XX
 PS Claim 20; Page 645-646; 765pp; English.
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU32655-2001-611725/70 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 CC
 SQ Sequence 2202 AA;
 Query Match 97.5%; Score 3783.5; DB 4; Length 2202;
 Best Local Similarity 84.3%; Pred. No. 3,4e-272;
 Matches 751; Conservative 0; Mismatches 1; Indels 139; Gaps 2;
 QY 1 MGFSSELSPQHGVLQOQWAELELLLEGKRMMAQVSDREYAGLHHMSLQDSGQS 60
 DB 1 MGFSSELSPQHGVLQOQWAELELLLEGKRMMAQVSDREYAGLHHMSLQDSGQS 60
 QY 61 RAISPDSPISQWABITQTEGLSRLRQHAEDLNSGLSKSLRLRERQQLRTYSEGM 120
 DB 61 RAISPDSPISQWABITQTEGLSRLRQHAEDLNSGLSKSLRLRERQQLRTYSEGM 120
 QY 121 QOLOELTTHQDIEKLKQYRALARDAQAKRYQESKQKDPKADKXVRSIMKLF 180
 DB 121 QOLOELTTHQDIEKLKQYRALARDAQAKRYQESKQKDPKADKXVRSIMKLF 180
 QY 181 AHNHRYVLGVRAAQLHHQHLLPGLRLSLQDLHEENACTIKETLQETLEISLVDDE 240
 DB 181 AHNHRYVLGVRAAQLHHQHLLPGLRLSLQDLHEENACTIKETLQETLEISLVDDE 240
 QY 241 VVAIHREMAAAARIQPEAEYQGFRLQYGSADVPVPCVTFDESLEEGEPLEGEQLQNE 300
 DB 241 VVAIHREMAAAARIQPEAEYQGFRLQYGSADVPVPCVTFDESLEEGEPLEGEQLQNE 300
 QY 301 LTVESVQHTLTSVTDELAVALATEMVFRRQEMVNTQLOELNNEENHTPRERVOQLGKROYL 360
 DB 301 LTVESVQHTLTSVTDELAVALATEMVFRRQEMVNTQLOELNNEENHTPRERVOQLGKROYL 360
 QY 361 QEALQGLQVALCSQAKLQAOQELQTKLEHLGPGEPPVLLQODDRHSTSSSQEREGGR 420
 DB 361 QEALQGLQVALCSQAKLQAOQELQTKLEHLGPGEPPVLLQODDRHSTSSSQEREGGR 420
 QY 421 TPTLEILKSHISGIFRPKFS----- 440
 DB 421 TPTLEILKSHISGIFRPKFS----- 440
 QY 441 -----NLVRLSEGGPSPILIDHLSTQOPLTKK 470
 DB 481 LVRESQKQBYVLVSLMDGLPRHFIITQSLDNLVRLSEGGPSPILIDHLSTQOPLTKK 540
 QY 471 SGVVLHRAVPKDKMVLNHEDLVLGEQIGRNGFGEVSGRLRADNTLVAVKSCHETLPPDL 530
 DB 541 SGVVLHRAVPKDKMVLNHEDLVLGEQIGRNGFGEVSGRLRADNTLVAVKSCHETLPPDL 600
 QY 531 KAKFLOBARILKQYSHPNVRLIGVCTQKOPITYINVELVOGGPPLFPLRTGEGARLVKTL 590
 DB 601 KAKFLOBARILKQYSHPNVRLIGVCTQKOPITYINVELVOGGPPLFPLRTGEGARLVKTL 660
 QY 591 LQWVGDAAGMEYLESKCCIHRLDAAARNCLVTEKNVLKISDFGMSREBAGVVAASGGRL 650
 DB 661 LQWVGDAAGMEYLESKCCIHRLDAAARNCLVTEKNVLKISDFGMSREBAGVVAASGGRL 720
 QY 651 QVPVKMTABEALNYGRYSSESDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRLLP 710
 DB 721 QVPVKMTABEALNYGRYSSESDVMSFGILLMETFSLGASPYPLSNQOTREFVEKGRLLP 780
 QY 711 CPRLCPDAVFRMLEQCMAYEPGQRPSPSTTYQELQSIIRKRR 752
 DB 781 CPRLCPDAVFRMLEQCMAYEPGQRPSPSTTYQELQSIIRKRR 822
 QY 822 -----GNFGEVFSGRRLRADNTLVAVKS 521

Db 723 CPELCPDAVFRMLEOCMAVEPGQRPSPFSTIYOELOSIRKRRH 764

RESULT 9
ADH43090
ID ADH43090 standard; protein; 820 AA.

XX
AC ADH43090;
XX
DT 25-MAR-2004 (first entry)

XX
DE GRAM related protein #SEQ ID 2.

XX
DE Neuroprotective; neutralisation; CRAM;
XX
KM collapsing-response mediator protein-associated molecule; gene therapy;
KM mitochondria; drug development; neural disease; pathosis.

XX
OS Mus musculus.

XX
PN WO2004001038-A1.

XX
PD 31-DEC-2003.

XX
PF 19-JUN-2003; 2003WO-JP007766.

XX
PR 19-JUN-2002; 2002JP-00179105.

XX
PA (NEW-) NEW IND RES ORG.

XX
PI Yanaqi S;

XX
DR N-PSDB; ADH43089.

XX
PT Genes and proteins participating in neutralisation of cells or tissues,
PT useful in gene therapy and regeneration medicine, applicable in
PT diagnosis, drug development for neural diseases and study of mechanism of
PT pathosis.

XX
PS Claim 2; SEQ ID NO 2; 101pp; Japanese.

XX
CC The invention relates to a method for inducing neutralisation of cells or
CC tissues by using a protein binding to CRAM (collapsing-response mediator
CC protein-associated molecule) protein or its encoded gene. The proteins
CC and their encoded genes are useful in gene therapy and regenerative
CC medicine, e.g. by inducing neutralisation of mitochondria. They are also
CC applicable in diagnosis, drug development for neural diseases and
CC studying the mechanism of pathosis. The current sequence represents a
CC CRAM related protein.

XX
SQ Sequence 820 AA;

Query Match 89.5%; Score 3473; DB 8; Length 820;
Best Local Similarity 82.4%; Pred. No. 1.2e-249;
Matches 677; Conservative 35; Mismatches 38; Indels 72; Gaps 2;

QY 1 MGFSSELCPQGHGVLOQMOEAEFLRLLEGMRKMAQRYKSDREYAGLLHHMSLODSGGOS 60
Db 1 MGFSSELCPQGHGVLOQMOEAEFLRLLEGMRKMAQRYKSDREYAGLLHHMSLODSGGOS 60

QY 61 RAISPDSPISQSWAETTSQTEGLRLLRQHAEDLNSGFLSKLILIRERQQLRTKYSBOM 120
Db 61 WSGGPPSPVQSWAETTSQTEGLRLLRQHAEDLNSGFLSKLILIRERISLAKTINEOK 120

QY 121 QQLQOEELTKTHSODIEKLKSYRALARDAQAQKRTQESKDYORDKADKTYRSLMKLP 180
Db 121 QQLQOEELTKTHSODIEKLKSYRALARDAQAQKRTQESKDYORDKADKTYRSLMKLP 180

QY 181 AHNHRYVLGVRAQOLHHQHHLHLLPGLRSLQDLHEHMACTIKELIQEYLETSSIVQDE 240
Db 181 AHNHRYVLGVRAQOLHHQHHLHLLPGLRSLQDLHEHMACTIKELIQEYLETSSIVQDE 240

QY 241 VVAIHREMAAAARIQPEAEYQGLRQYGSAPDVPCTVDESLLEGGEPLBEGELQUNE 300

Db 241 VASIHREMAAAARIQPEAEYQGLRQYGSAPDVPCTVDESLLEGGEPLBEGELQUNE 300

QY 301 LTVESVQHTLTSYDEELAVATEMVFRRQEMTQLOQELRNEENHPRRVOQLKROYL 360
Db 301 LTVESVQHTLTSYDEELAVATEMVFRRQEMTQLOQELRNEENHPRRVOQLKROYL 360

QY 361 DEALQGLVALCSQAKLQAOQELLOTLEHLGPGEPPLYLQDDRHSSTSSSEOREGGR 420
Db 361 DEALQGLVALCSQAKLQAOQELLOTLEHLGPGEPPLYLQDDRHSSTSSSEOREGGR 420

QY 421 TPTLEILKSHISGIRPKRS----- 440
Db 421 TPTLEILKSHISGIRPKRS----- 440

QY 441 TPTLEILKSHISGIRPKRS----- 440
Db 441 TPTLEILKSHISGIRPKRS----- 440

QY 479 LVRESQGRQEVYLVSMWMDGPHRHFILOSUDNLYRLEGDFPSIPLILTHLSSQOPLTK 538
Db 479 LVRESQGRQEVYLVSMWMDGPHRHFILOSUDNLYRLEGDFPSIPLILTHLSSQOPLTK 538

QY 471 SGVTLHRAVPRKQVNLNHEDLVLAGEQLGRNGEYFSGRLRADNTLVAVKSCRETLPPDL 530
Db 471 SGVTLHRAVPRKQVNLNHEDLVLAGEQLGRNGEYFSGRLRADNTLVAVKSCRETLPPDL 530

QY 539 SGVTLHRAVPRKQVNLNHEDLVLAGEQLGRNGEYFSGRLRADNTLVAVKSCRETLPPDL 598
Db 539 SGVTLHRAVPRKQVNLNHEDLVLAGEQLGRNGEYFSGRLRADNTLVAVKSCRETLPPDL 598

QY 531 KAKFLQEARILIKQYHPNIVRLIGVCTQKOPYYIMELVQSGDFLFTLBEGARLRVKT 590
Db 531 KAKFLQEARILIKQYHPNIVRLIGVCTQKOPYYIMELVQSGDFLFTLBEGARLRVKT 590

QY 599 KAKFLQEARILIKQYHPNIVRLIGVCTQKOPYYIMELVQSGDFLFTLBEGARLRVKT 658
Db 599 KAKFLQEARILIKQYHPNIVRLIGVCTQKOPYYIMELVQSGDFLFTLBEGARLRVKT 658

QY 591 LQWGDAAAGMEYLESKCCIRHDLAARNCILTEKXVNLKISDGMGRBEADGYAASGGLR 650
Db 591 LQWGDAAAGMEYLESKCCIRHDLAARNCILTEKXVNLKISDGMGRBEADGYAASGGLR 650

QY 659 LQWGDAAAGMEYLESKCCIRHDLAARNCILTEKXVNLKISDGMGRBEADGYAASGGLR 718
Db 659 LQWGDAAAGMEYLESKCCIRHDLAARNCILTEKXVNLKISDGMGRBEADGYAASGGLR 718

QY 651 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSLGASPYNLSNOQTRFEVKGGRLP 710
Db 651 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSLGASPYNLSNOQTRFEVKGGRLP 710

QY 719 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSLGASPYNLSNOQTRFEVKGGRLP 778
Db 719 QVPVKMTAPEALNYGRYSSESQVMSFGILLMETFSLGASPYNLSNOQTRFEVKGGRLP 778

QY 711 CPELCPDAVFRMLEOCMAVEPGQRPSPFSTIYOELOSIRKRRH 752
Db 711 CPELCPDAVFRMLEOCMAVEPGQRPSPFSTIYOELOSIRKRRH 752

QY 779 CPELCPDAVFRMLEOCMAVEPGQRPSPFSTIYOELOSIRKRRH 820
Db 779 CPELCPDAVFRMLEOCMAVEPGQRPSPFSTIYOELOSIRKRRH 820

RESULT 10
AAU32654
ID AAU32654 standard; protein; 2017 AA.

XX
AC AAU32654;
XX
DT 18-DEC-2001 (first entry)

XX
DE Novel human secreted protein #3145.

XX
KM Human; vaccination; gene therapy; nutritional supplement;
KM stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KM immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX
OS Homo sapiens.

XX
PN WO200179449-A2.

XX
PD 25-OCT-2001.

XX
PF 16-APR-2001; 2001WO-US008656.

XX
PR 18-APR-2000; 2000US-00552929.

XX
PR 26-JAN-2001; 2001US-00770160.

XX
PA (HYSE-) HYSBO INC.

XX
PI Tang YT, Liu C, Drmanac RT;

XX
DR WPI, 2001-611725/70.

XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.

XX Claim 20, Page 645; 765pp; English.

CC The invention relates to novel human secreted polypeptides: The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention

XX Sequence 2017 AA;

Query Match 62.0%; Score 2405; DB 4; Length 2017;
 Best Local Similarity 73.1%; Pred. No. 1,6e-169;
 Matches 488; Conservative 0; Mismatches 0; Indels 180; Gaps 3;

QY 224 KEILQYELISSIVQDEVAIHEMAAAARIQPEAEYQGLRQYGSADVPVCTFDES 283
 DB 1 KEILQYELISSIVQDEVAIHEMAAAARIQPEAEYQGLRQYGSADVPVCTFDES 60
 QY 284 LIEEGPRLPEGLQNLNLTVESVQHTLTSVTDELAIVATEVPRRQSMVTOLOQLNNEE 343
 DB LIEEGPRLPEGLQNLNLTVESVQHTLTSVTDELAIVATEVPRRQSMVTOLOQLNNEE 86
 QY 344 NTHPRERVOLLGKRYQALQGLQVALCSQAKLQAOQELLQTKLEHLPGEPPVLLQ 403
 DB 87 VOLLGKRYQALQGLQVALCSQAKLQAOQELLQTKLEHLPGEPPVLLQ 139
 QY 404 DDRHSTSSSEQEREGRTPTLEILKSHIGIFRPKFS----- 440
 DB 140 DDRHSTSSSEQEREGRTPTLEILKSHIGIFRPKFSLPPLQLIPEVQKPLHEQLWYHG 199
 QY 441 ----- 440
 DB 200 AIPRAEVAELVHSGDFLVRESQKQEVYLSTVLMQGLPRHFIIOQLDGSRLPMEADPG 259
 QY 441 -----NLVRLLEGSGFPISILLIDHLLSTOQPLTKSGVVLHRAVPRKXVLANHEDLVLE 495
 DB 260 SPALQNLVRLLEGSGFPISILLIDHLLSTOQPLTKSGVVLHRAVPRKXVLANHEDLVLE 319
 QY 496 QIGR-----GNFGE 504
 DB 320 QIGRVPRQSGNSGRVAVRGNTGAPHPGVGSRWGRKRRRLRWMEGRGSPRPFQGNFGE 379
 QY 505 VPSGRRLRANNTLVAVKSCRETLPPDLKAKFLQEPARILKQYSHNRYRLIGVCTQOKPIYI 564
 DB 380 VPSGRRLRANNTLVAVKSCRETLPPDLKAKFLQEPARILKQYSHNRYRLIGVCTQOKPIYI 439
 QY 565 VNELVGGDFLPTLRTEGARLRVKTLLQWGDAAAGMEVLESCTCHRDLAANCLVTER 624
 DB 440 VNELVGGDFLPTLRTEGARLRVKTLLQWGDAAAGMEVLESCTCHRDLAANCLVTER 499
 QY 625 NVLKISDFGMSREADGVYAAASGLRQVPVKWTAPALNVGRYSSESDVWSPFILLMETF 684
 DB 500 NVLKISDFGMSREADGVYAAASGLRQVPVKWTAPALNVGRYSSESDVWSPFILLMETF 559
 QY 685 SIGASYPNLSNQCTEPEVEKGRLLPCPELCPAVALRLMEQCAAYEPGQRPSTTYOEL 744
 DB 560 SIGASYPNLSNQCTEPEVEKGRLLPCPELCPAVALRLMEQCAAYEPGQRPSTTYOEL 619
 QY 745 QSIRKRRH 752
 DB 745 QSIRKRRH 752

DB 620 QSIRKRRH 627

RESULT 11
 ADK71829
 ID ADK71829 standard; protein; 472 AA.
 XX
 AC ADK71829;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Human kinase and phosphatase KPP-6 protein.
 XX
 KW human; kinase; phosphatase; KPP; cardiovascular; antiarteriosclerotic;
 KW hypotensive; vasotropic; antiinflammatory; antianginal; anti-HIV;
 KW antiallergic; antiasthmatic; immunosuppressive; antithyroid;
 KW dermatological; antidiabetic; nephrotropic; angiot; gastrointestinal;
 KW neuroprotective; osteopathic; antiarthritic; uropathic; ophthalmological;
 KW antirheumatic; antiparkinsonian; nootropic; anticonvulsant; hepatotropic;
 KW antipneumatic; haemostatic; cytostatic; antilipemic; antiparasitic;
 KW antihelminthic; antibacterial; virucide; protozoacide; fungicide;
 KW cardiovascular disease; immune system; neurological; growth; development;
 KW cell proliferation; viral; bacterial; fungal; parasitic; protozoan;
 KW helminthic infection; transgenic; gene therapy; enzyme;
 KW single nucleotide polymorphism; SNP.
 XX
 OS Homo sapiens.
 PN WO2004018641-A2.
 XX
 PD 04-MAR-2004.
 XX
 PF 25-AUG-2003; 2003WO-US026635.
 XX
 PR 26-AUG-2002; 2002US-0406172P.
 PR 25-SEP-2002; 2002US-0413910P.
 PR 27-SEP-2002; 2002US-0414296P.
 PR 11-OCT-2002; 2002US-0417821P.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Baughn MR, Richardson TW, Marguis JP, Swarnakar A, Tang YT;
 PI Becha SD, Emerling BM, Jin P, Wilson AD, Yue H, Gietzen KJ;
 PI Chang H, Yang YG, Lee SY, Khare R, Elliott VS, Hatalla AJA;
 PI Chawla NK, Ramkumar J, Gururajan R, Tribouley CM, Chien D, Tran UK;
 PI Murgue J;
 XX
 DR WPI: 2004-226830/21.
 DR N-PSDB; ADK71888.
 XX
 PT New human kinases and phosphatases, useful for diagnosing, treating or
 PT preventing atherosclerosis, hypertension, AIDS, allergy, multiple
 PT sclerosis, osteoarthritis, Alzheimer's disease, Crohn's disease, cancer
 PT or hepatitis.
 XX
 PS Claim 1; SEQ ID NO 6; 347pp; English.
 XX
 CC The invention relates to a novel isolated polypeptide which is a human
 CC kinase and phosphatase (KPP). The polypeptide of the invention
 CC demonstrates cardiovascular, antiarteriosclerotic, hypotensive,
 CC vasotropic, antiinflammatory, antianginal, anti-HIV, antiallergic,
 CC antiasthmatic, immunosuppressive, antithyroid, dermatological,
 CC osteopathic, nephrotropic, angiot, gastrointestinal, neuroprotective,
 CC antipneumatic, antiparkinsonian, nootropic, anticonvulsant, hepatotropic,
 CC haemostatic, cytostatic, antilipemic, antiparasitic, antihelminthic,
 CC antibacterial, virucide, protozoacide and fungicide activities. The
 CC kinase and phosphatase (KPP) polynucleotides, polypeptides, agonists and
 CC antagonists may be useful for diagnosing, treating or preventing
 CC disorders such as cardiovascular diseases, immune system disorders,
 CC neurological disorders, disorders affecting growth and development, cell
 CC proliferative disorders and viral, bacterial, fungal, parasitic, cell
 CC protozoan or helminthic infections. Furthermore, the molecules of the

DB 774 EGGYMSAPQHCPEDISKIMKCMWDYKPEBNRPKFSLEQELTIIKK 820

RESULT 13
ADL71053
ID ADL71053 standard; protein; 823 AA.
XX
AC ADL71053;
XX
DT 20-MAY-2004 (first entry)
XX
XX
DE Type II collagen expression promoting protein, seq id 46.
XX
KM Osteopathic; antiinflammatory; antiarthritis; antiarthritic;
KM gene therapy; type II collagen; expression; cartilage disease;
KW osteoarthritis; cartilage defect; rheumatoid arthritis; human.
XX
OS Homo sapiens.
XX
PN WO2003087375-A1.
XX
PD 23-OCT-2003.
XX
PF 16-APR-2003; 2003WO-JP004802.
XX
PR 16-APR-2002; 2002JP-00113908.
PR 19-APR-2002; 2002US-0373594P.
XX
PA (ASAH) ASAMI KASEI KK.
XX
PI Matsuda A, Honda G, Muramatsu S;
XX
DR WPI; 2003-845331/78.
DR N-PSDB; ADL71052.
XX
XX
PT New purified protein that promotes type II collagen expression, useful
PT for preventing and treating a cartilage disease, e.g. osteoarthritis,
PT cartilage defect, or rheumatoid arthritis.
XX
PS Claim 1; SEQ ID NO 46; 271pp; English.
XX
XX
SQ Sequence 823 AA;
XX

Query Match 47.1%; Score 1826.5; DB 7; Length 823;
Best Local Similarity 45.5%; Pred. No 5.8e-127;
Matches 376; Conservative 147; Mismatches 219; Indels 85; Gaps 10;

QY 1 MGFSSSELCPQGHGVLQOQAEELRLLEGKMKMAQVXSDBREYAGLHMSIQDSGGQS 60
DB 1 MGFSGSLDKNSQ-EAVL-KLQDMELRLLETVKKFMALRIKSDKEVATLQNLGNQVDEKST 58
QY 61 RAISPPSPISQSWAETTSQTEGSLRLRQHAEDLNSGPIKSLILREKQQLAKTYSBOW 120
DB 59 VQVNVYSNVSKGMLMIQOTEQSLRIKTHAEDLNSGPIRLTMIKDQVVKKSYVGIN 118
QY 121 OOLQOELTYTHSODIEKLKSOYRALARDAQAQKTKQEA-SKQDDBKAKDKVVRSLMTKL 179
DB 119 QQIEBAMIKVTETELKELKSSYRQLIKEMNSAKETKEALAKGETEKAKERTDKATMKL 178
QY 180 FAHHRNYVLGVAQAOLHONHQLLPGLRLSLQDLHEEMACTLKELQEYLEISSLYOD 239
DB 179 HMLHNGYVALAKGAQHOSQYVDTTLPLLDSDVQKQOEEMIKALKGIFPDYSGITSLVME 238
QY 240 EYVAIYHREMAAAAAARIQPEAYIGFURQYGSAPDVPCTTFDESLLEBGEPLERGLQIN 299

DB 239 EIVNVHKEIQMSVEQIDPESTETNNPFDVHRTIAAKGELEFDTSLSEENENLOANEIMWN 298
QY 300 ELTVESVOHTLTSVTDDELAVATEWVFRQEWNTOLQOELRNEENTHPR-ERVOLLGKQ 358
DB 299 NLTADSLQVMLKTLABELVTOQOMLHKEAAVLBLEKRLEESFECEKSDYLLGLGQK 358
QY 359 VLOEALQGLQVNLCSQAKLQAOQELQTLERHLEGEPEPPVULLQDDRRSTSSSEOREG 418
DB 359 ALEELKQSVQOARCEAKCAQKALLEQKVOENDGKEPPVNVYBEDARSVTSMEKE-- 416
QY 419 GRPTLEIKSHISGIFR-PK----- 438
DB 417 -RLSKFESIRSHIAGIKSPKSVLGSSTQVCVIVSGERPLAHDWYHGAIRIENQELL 475
QY 439 -----FSNLYRLBGGPSPILIDHLLST 463
DB 476 KQGGDFLVNESHGKPEEYVLVYSDQQRHPIIQFVDNLVRFEGTGFSNIPOLIDHFMFT 535
QY 464 QOPLTKSGVYLHRAVPKD-KVYLNHEDLVLEQIGRGKFGVPSGRLADNTLVAVVKSC 522
DB 536 KQVITTKSGVVLNPIPKDKKWLNHEDVSLGELLGKNGFGEVYKGTLK-DKTPVALIKTC 594
QY 523 RETLPEDLKAKFLQEARILKQYSHPNIVRLIGVCTOKOPIYIYIMELVQSGDPLTFRTG 582
DB 595 KEDLPQELKIKFLQEKILKQYDHNIVKLVGVCTORQRYIIMELVPGDPLTFIRKX 654
QY 583 ARLRVYTLQVNGDAAGMEYLESKQCIHRDLAARNCVTEKQVNLKISDFGMSREBAGV 642
DB 655 DELKDKQVRFSLDVAAGMLYLESKNCIHRDLAARNCVGENNTLKISDFGMSRQEDGV 714
QY 643 YAAAGGLQVYVYKWTAPPEALNAGRYSSBVDVMSFGILMETSLGASPIPNLSNOGTREF 702
DB 715 YSSS-GLKQIPIKWAPPEALNAGRYSSBVDVMSFGILMETPSLGCPYPGMTNOQARQ 773
QY 703 VKKGRLLPCELPDPAVFRIMECWAHYEBCQRPSPFTIVQELQSIK 749
DB 774 VERGTRMSAPQNCPESEVFTIMKCMWDYKPEBNRPKFNIDLKELTVIYK 820

RESULT 14
AA43963
ID AA43963 standard; protein; 262 AA.
XX
AC AA43963;
XX
DT 21-DEC-1999 (first entry)
XX
DE Human protein kinase #18.
XX
KM Prediction; secondary structure; alignment; evolutionary conservation;
KW homology; periodicity; co-variation analysis; antigenic site;
XX site directed mutagenesis; interaction.
XX
OS Homo sapiens.
XX
PN US958784-A.
XX
PD 28-SEP-1999.
XX
PF 25-MAR-1992; 92US-00857224.
XX
PR 25-MAR-1992; 92US-00857224.
XX
PA (BENNER) BENNER S A.
XX
PI Benner SA;
XX
DR WPI; 1999-570766/48.
XX
PT Predicting the folded structure of proteins.
XX
PS Disclosure; Col 271-274; 113pp; English.

XX Sequences AAY43902-Y44015 represent proteins used in a novel method of
CC predicting the folded structure of proteins, by aligning sequences of
CC homologous proteins and using patterns of evolutionarily conserved and
CC varied sequences to assign positions. Positions in the alignment are
CC assigned to the surface or inside of the folded structure, active sites,
CC and pairing segments. Secondary structural units are assigned by
CC identifying periodicity in the assignments, and assembled into global
CC form using distance constraints imposed by disulfide bridges, active site
CC assignments and co-variation analysis. The predicted secondary structures
CC are useful for identifying antigenic sites on a protein molecule, as
CC guides for site directed mutagenesis studies, and for understanding the
CC interaction of a protein with other molecules

XX Sequence 262 AA;

Query Match 35.2%; Score 1364.5; DB 2; Length 262;
Best Local Similarity 99.6%; Pred. No. 3.2e-93;
Matches 262; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

DB 489 EDLVIGEQTGRNFGVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 548
1 EDLVIGEQTGRNFGVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 60
QY 549 IYRLIGVCTQKOPRIYIMELVGGDFLFLRTGRLRVKTLLOWVGDAAAGMEYLESKC 608
61 IYRLIGVCTQKOPRIYIMELVGGDFLFLRTGRLRVKTLLOWVGDAAAGMEYLESKC 120
DB 609 CIHRDLAARNCLVTEKNVLKISDFGMSREADGVYAASGLRQVPVKWTAPEALNYGRYS 668
121 CIHRDLAARNCLVTEKNVLKISDFGMSREADGVYAASGLRQVPVKWTAPEALNYGRYS 179
QY 669 SESDWSFGILLMETFSLGASPYPNLSNOOTREFVEKGRLPCEPLCPDAVRLMEQCWA 728
DB 180 SESDWSFGILLMETFSLGASPYPNLSNOOTREFVEKGRLPCEPLCPDAVRLMEQCWA 239
QY 729 YEPGQRPSPFTIYQELQSIKRRH 751
DB 240 YEPGQRPSPFTIYQELQSIKRRH 262

RESULT 15

AAY43964 standard; protein; 262 AA.

XX AAY43964;

DT 17-OCT-2003 (revised)
DT 21-DEC-1999 (first entry)

DE Cat protein kinase #1.

KM Prediction; secondary structure; alignment; evolutionary conservation;
KW homology; periodicity; co-variation analysis; antigenic site;
KW site directed mutagenesis; interaction.

OS Fells catus.

PN US5958784-A.

PD 28-SEP-1999.

PF 25-MAR-1992; 92US-00857224.

PR 25-MAR-1992; 92US-00857224.

PA (BENN/) BENNER S A.

PI Benner SA;

DR WPI, 1999-570766/48.

PT Predicting the folded structure of proteins.

XX PS Disclosure; Col 275-276; 113pp; English.

XX Sequences AAY43902-Y44015 represent proteins used in a novel method of
CC predicting the folded structure of proteins, by aligning sequences of
CC homologous proteins and using patterns of evolutionarily conserved and
CC varied sequences to assign positions. Positions in the alignment are
CC assigned to the surface or inside of the folded structure, active sites,
CC and pairing segments. Secondary structural units are assigned by
CC identifying periodicity in the assignments, and assembled into global
CC form using distance constraints imposed by disulfide bridges, active site
CC assignments and co-variation analysis. The predicted secondary structures
CC are useful for identifying antigenic sites on a protein molecule, as
CC guides for site directed mutagenesis studies, and for understanding the
CC interaction of a protein with other molecules. (Updated on 17-OCT-2003 to
CC standardise OS field)

XX Sequence 262 AA;

Query Match 34.8%; Score 1350.5; DB 2; Length 262;
Best Local Similarity 97.7%; Pred. No. 3.5e-92;
Matches 257; Conservative 4; Mismatches 1; Indels 1; Gaps 1;

QY 489 EDLVIGEQTGRNFGVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 548
1 EDLVIGEQTGRNFGVSGRLRADNTLVAVKSCRETLPPDLKAKFLQEARILKQYSHN 60
DB 549 IYRLIGVCTQKOPRIYIMELVGGDFLFLRTGRLRVKTLLOWVGDAAAGMEYLESKC 608
61 IYRLIGVCTQKOPRIYIMELVGGDFLFLRTGRLRVKTLLOWVGDAAAGMEYLESKC 120
QY 609 CIHRDLAARNCLVTEKNVLKISDFGMSREADGVYAASGLRQVPVKWTAPEALNYGRYS 668
DB 121 CIHRDLAARNCLVTEKNVLKISDFGMSREADGVYAASGLRQVPVKWTAPEALNYGRYS 179
QY 669 SESDWSFGILLMETFSLGASPYPNLSNOOTREFVEKGRLPCEPLCPDAVRLMEQCWA 728
DB 180 SESDWSFGILLMETFSLGASPYPNLSNOOTREFVEKGRLPCEPLCPDAVRLMEQCWA 239
QY 729 YEPGQRPSPFTIYQELQSIKRRH 751
DB 240 YEPGQRPSPFTIYQELQSIKRRH 262

Search completed: March 18, 2005, 15:59:19
Job time : 176 secs

Db 192 GACAGGAGATGATGACGAGCTGTTCAACCAATGTCCTGACGAGCAGTGGGGCCAGAGC 251
Qy 181 CGGGCCATCAGCCCTGACAGCCCATCAGTCACTCTGGGCTGAGATCACAGCCAACT 240
Db 252 CGGGCCATCAGCCCTGACAGCCCATCAGTCACTCTGGGCTGAGATCACAGCCAACT 311
Qy 241 GAGGGCTGAGCCCTGCTGCTGCGGACGACGAGAGGATCTGAATCAGGGCCCTGAGC 300
Db 312 GAGGGCTGAGCCCTGCTGCTGCGGACGACGAGAGGATCTGAATCAGGGCCCTGAGC 371
Qy 301 AAGCTGAGCTGCTCATCCGGGAAACGGGCAAGCTTGGCAAGACTTACAGCGAGCATG 360
Db 372 AAGCTGAGCTGCTCATCCGGGAAACGGGCAAGCTTGGCAAGACTTACAGCGAGCATG 431
Qy 361 CAGCAGCTGAGAGAGAGCTCAACAGAGCCCAAGCCAGAGCAATTGAGAGCTGAAGAGC 420
Db 432 CAGCAGCTGAGAGAGAGCTCAACAGAGCCCAAGCCAGAGCAATTGAGAGCTGAAGAGC 491
Qy 421 CAGTACCGAGCTGAGAGAGCTGAGAGCCAGAGCCAGAGCAAGTACAGAGAGCCAGC 480
Db 492 CAGTACCGAGCTGAGAGAGCTGAGAGCCAGAGCCAGAGCAAGTACAGAGAGCCAGC 551
Qy 481 AAGAGCAGAGAGCCGTGACAAAGCCAAAGAGCAAGTATGTGCGAGAGCTGTGAAGCTTT 540
Db 552 AAGAGCAGAGAGCCGTGACAAAGCCAAAGAGCAAGTATGTGCGAGAGCTGTGAAGCTTT 611
Qy 541 GCTCAACACACCGCTATGTGCTGGGGGTGGGGCTGGGAGCTTACACCAAGCCAGCCAGC 600
Db 612 GCTCAACACACCGCTATGTGCTGGGGGTGGGGCTGGGAGCTTACACCAAGCCAGCCAGC 671
Qy 601 CACGAGCTCTGCTGAGCCGAGCTGCTGCGGTCACTGAGAGAGCTGACGAGAGATGAGT 660
Db 672 CACGAGCTCTGCTGAGCCGAGCTGCTGCGGTCACTGAGAGAGCTGACGAGAGATGAGT 731
Qy 661 TGCATCTGAAGAGAGATCTGTCAGAGAAATACCTGAGAAATAGCAAGCTGTGTCAGATGAG 720
Db 732 TGCATCTGAAGAGAGATCTGTCAGAGAAATACCTGAGAAATAGCAAGCTGTGTCAGATGAG 791
Qy 721 GTGGTGGCCATTCAACCGGAGATGAGCTGAGAGCTGCGCCGCAATCCAGCCCTGAGGCTGAG 780
Db 792 GTGGTGGCCATTCAACCGGAGATGAGCTGAGAGCTGCGCCGCAATCCAGCCCTGAGGCTGAG 851
Qy 781 TACCAAGGCTTCCGTCGACAGTATGAGGTCCGACCTGACCTCCACCTGTGTCAAGTTTC 840
Db 852 TACCAAGGCTTCCGTCGACAGTATGAGGTCCGACCTGACCTCCACCTGTGTCAAGTTTC 911
Qy 841 GATGAGTCACTGCTTGAAGAGAGGTGAACGCTGAGAGCTTGGAGAGCTTCAAGTGAAGAG 900
Db 912 GATGAGTCACTGCTTGAAGAGAGGTGAACGCTGAGAGCTTGGAGAGCTTCAAGTGAAGAG 971
Qy 901 CTGAGCTGAGAGAGCTGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGC 960
Db 972 CTGAGCTGAGAGAGCTGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGC 1031
Qy 961 ACCGAGATGAGTGTTCAGAGGCGGACAGAGATGATGATGACAGCTGCAACAGAGACTCCGGAAT 1020
Db 1032 ACCGAGATGAGTGTTCAGAGGCGGACAGAGATGATGATGACAGCTGCAACAGAGACTCCGGAAT 1091
Qy 1021 GAAAGAGAGAAACCCACCCCGGAGAGCGGAGTGAAGCTGCTGAGAGAGAGAGAGAGAGTCTG 1080
Db 1092 GAAAGAGAGAAACCCACCCCGGAGAGCGGAGTGAAGCTGCTGAGAGAGAGAGAGAGTCTG 1151
Qy 1081 CAAGAGAGACTGACAGAGAGAGCTGAGAGTACGCTGTGACAGAGAGAGAGAGAGAGAGAG 1140
Db 1152 CAAGAGAGACTGACAGAGAGAGCTGAGAGTACGCTGTGACAGAGAGAGAGAGAGAGAGAG 1211
Qy 1141 CAGAGATGCTGACAGAGAGAGCTGAGAGTACGCTGTGACAGAGAGAGAGAGAGAGAGAG 1200
Db 1212 CAGAGATGCTGACAGAGAGAGCTGAGAGTACGCTGTGACAGAGAGAGAGAGAGAGAGAG 1271
Qy 1201 CTCTCTGAGAGATGACAGAGAGCTGAGAGTACGCTGTGACAGAGAGAGAGAGAGAGAG 1260
Db 1272 CTCTCTGAGAGATGACAGAGAGCTGAGAGTACGCTGTGACAGAGAGAGAGAGAGAGAG 1331

Qy 1261 AACCCAGAGCTGAGAGATCTTAAAGAGCCATCTCAGAGATCTTCCGCCCCAAAGTTCTCG 1320
Db 1332 AACCCAGAGCTGAGAGATCTTAAAGAGCCATCTCAGAGATCTTCCGCCCCAAAGTTCTCG 1391
Qy 1321 AACCTGTACCGACTGAGAGAGAGAGCTTCTCTAGAGATCTTCTGCTCATGACACCTTA 1380
Db 1392 AACCTGTACCGACTGAGAGAGAGAGCTTCTCTAGAGATCTTCTGCTCATGACACCTTA 1451
Qy 1381 CTGAG 1440
Db 1452 CTGAG 1511
Qy 1441 AAG 1500
Db 1512 AAG 1571
Qy 1501 AACCTTGGAGAGAGTGTTCAGCGGAGCGCTGCGAGAGCCGACCAACACCTGTGTGGGTGAAG 1560
Db 1572 AACCTTGGAGAGAGTGTTCAGCGGAGCGCTGCGAGAGCCGACCAACACCTGTGTGGGTGAAG 1631
Qy 1561 TCTTGTGAGAGAGAGCTCCCACTGACCTCAAGAGCCAAATTTCTTAAGAGAGAGAGATC 1620
Db 1632 TCTTGTGAGAGAGAGCTCCCACTGACCTCAAGAGCCAAATTTCTTAAGAGAGAGAGATC 1691
Qy 1621 CTGAAGCAGTACAG 1680
Db 1692 CTGAAGCAGTACAG 1751
Qy 1681 CCCATCTACATGCTGATGAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1752 CCCATCTACATGCTGATGAGAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1811
Qy 1741 GAGGG 1800
Db 1812 GAGGG 1871
Qy 1801 ATGAGATTAATGAG 1860
Db 1872 ATGAGATTAATGAG 1931
Qy 1861 GTGACAGAGAGAGATGCTCTGAAGATCACTTGTGAGATGTCCTGAGAGAGAGAGAGAG 1920
Db 1932 GTGACAGAGAGAGATGCTCTGAAGATCACTTGTGAGATGTCCTGAGAGAGAGAGAGAG 1991
Qy 1921 GGGGTATATGAG 1980
Db 1992 GGGGTATATGAG 2051
Qy 1981 GGCCTTAATCAAG 2040
Db 2052 GGCCTTAATCAAG 2111
Qy 2041 TGGAGAGAGCTTCAAGCTGAG 2100
Db 2112 TGGAGAGAGCTTCAAGCTGAG 2171
Qy 2101 GAGTTGTGAG 2160
Db 2172 GAGTTGTGAG 2231
Qy 2161 AGGCTCATGAG 2220
Db 2232 AGGCTCATGAG 2291
Qy 2221 TACCAAG 2256
Db 2292 TACCAAG 2327

RESULT 2
us-10-660-763-1
; Sequence 1, Application US/10660763

Publication No. US20040063130A1
GENERAL INFORMATION:
APPLICANT: GAN, Weiniu et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001183DIV1
CURRENT APPLICATION NUMBER: US/10/660,763
CURRENT FILING DATE: 2003-09-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2674
TYPE: DNA
ORGANISM: Homo sapiens
US-10-660-763-1

Query Match 100.0%; Score 2256; DB 17; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCTTCTCTCTGAGCTGTGAGCCGCCAGGAGCCAGGGGTCTTGACAGCAATGACAG 60
DB 72 ATGGGCTTCTCTCTGAGCTGTGAGCCGCCAGGAGCCAGGGGTCTTGACAGCAATGACAG 131
QY 61 GAGGCGGAGCTTGTCTTACTGAGAGGCGATGAAAGTGAATGGCCAGCGGTCTGAGAGT 120
DB 132 GAGGCGGAGCTTGTCTTACTGAGAGGCGATGAAAGTGAATGGCCAGCGGTCTGAGAGT 191
QY 121 GACAGGAGATGATCAGAGAGCTTCTACCATGTCCTTGAGAGCAGTGGGGCCAGAGC 180
DB 192 GACAGGAGATGATGAGAGCTTCTACCATGTCCTTGAGAGCAGTGGGGCCAGAGC 251
QY 181 CGGAGCATCAGCCCTGACAGCCCATCATGTCAGTCTGGGCTGAGATCACAGCCAACT 240
DB 252 CGGAGCATCAGCCCTGACAGCCCATCATGTCAGTCTGGGCTGAGATCACAGCCAACT 311
QY 241 GAGGCGCTGAGCGCTTGTCTGCGGAGCAGCAGAGAGATTGAATCTGAGGGCCCTGAGC 300
DB 312 GAGGCGCTGAGCGCTTGTCTGCGGAGCAGCAGAGAGATTGAATCTGAGGGCCCTGAGC 371
QY 301 AAGCTGAGCTGCTCATTCGGGAAACGGGACAGCTTGGCAAGCTTACAGGAGAGCTGG 360
DB 372 AAGCTGAGCTGCTCATTCGGGAAACGGGACAGCTTGGCAAGCTTACAGGAGAGCTGG 431
QY 361 CAGCAGCTGAGCAGAGCTCACCAAGACCCAGCCAGCAGCAGCATTGAGAGCTGAGAGC 420
DB 432 CAGCAGCTGAGCAGAGCTCACCAAGACCCAGCCAGCAGCAGCATTGAGAGCTGAGAGC 491
QY 421 CAGTACCGAGCTCTGGCAGCGGACAGTGGCCAAAGCCAAAGTACCAAGAGCCAGC 480
DB 492 CAGTACCGAGCTCTGGCAGCGGACAGTGGCCAAAGCCAAAGTACCAAGAGCCAGC 551
QY 481 AAGAGCAAGGACCGTGAACAAGGACCAAGATATGTGGCAGCCGTGGGAAGCTTTT 540
DB 552 AAGAGCAAGGACCGTGAACAAGGACCAAGATATGTGGCAGCCGTGGGAAGCTTTT 611
QY 541 GCTCACCAACAACGCTATGTGTGGGCGTGGGAGCTGCGAGCTACACACACACAC 600
DB 612 GCTCACCAACAACGCTATGTGTGGGCGTGGGAGCTGCGAGCTACACACACACAC 671
QY 601 CACACAGCTCTGCTGCGCGGCTGCTGCGGTCACTGAGAGACCTGCAAGAGAGTGGCT 660
DB 672 CACACAGCTCTGCTGCGCGGCTGCTGCGGTCACTGAGAGACCTGCAAGAGATGGCT 731
QY 661 TGATCTGAAGAGATCTGCAAGAAATACCTGGAATTGAGAGCCGTGGAGAGTGG 720
DB 732 TGATCTGAAGAGATCTGCAAGAAATACCTGGAATTGAGAGCCGTGGAGAGTGG 791
QY 721 GTGTGGCATTGACCGGAGATGGCTGAGCTGTCCCGCATCCAGCCTGAGGCTGAG 780
DB 792 GTGTGGCATTGACCGGAGATGGCTGAGCTGTCCCGCATCCAGCCTGAGGCTGAG 851

QY 781 TACCAAGCTTCTCTGAGCAGTATGGTCCGACCTGACGTCCACCCCTGTGTACAGTTC 840
DB 852 TACCAAGCTTCTCTGAGCAGTATGGTCCGACCTGACGTCCACCCCTGTGTACAGTTC 911
QY 841 GATGAGTCACTGCTTGAAGAGGATGAGACCTGAGACCTGGGAGCTGCACTGAGAG 900
DB 912 GATGAGTCACTGCTTGAAGAGGATGAGACCTGAGACCTGGGAGCTGCACTGAGAG 971
QY 901 CTGACTGTGAGAGCGTGAAGCAGACGCTGACCTCAGTGAACAGATGAGCTGTGGCC 960
DB 972 CTGACTGTGAGAGCGTGAAGCAGACGCTGACCTCAGTGAACAGATGAGCTGTGGCC 1031
QY 961 ACCGAGATGTGTTCAGGCGGCGAGGAGATGTTAACGACTGTGAACAGAGCTCCGGAAT 1020
DB 1032 ACCGAGATGTGTTCAGGCGGCGAGGAGATGTTAACGACTGTGAACAGAGCTCCGGAAT 1091
QY 1021 GAAGAGAGAACACCCACCCCGGAGCGGAGTGAAGCTGCTGGGCAAGAGCAAGTCTG 1080
DB 1092 GAAGAGAGAACACCCACCCCGGAGCGGAGTGAAGCTGCTGGGCAAGAGCAAGTCTG 1151
QY 1081 CAAGAGCACTGAGGAGGCTGAGAGTGAAGCTGTGAGCAGGCAAGCTGACGCGCCAG 1140
DB 1152 CAAGAGCACTGAGGAGGCTGAGAGTGAAGCTGTGAGCAGGCAAGCTGACGCGCCAG 1211
QY 1141 CAGGAGTTGTGAGAACCAAGCTGAGACCTGTGGGCGCGGAGAGCCCGCTGTGTG 1200
DB 1212 CAGGAGTTGTGAGAACCAAGCTGAGACCTGTGGGCGCGGAGAGCCCGCTGTGTG 1271
QY 1201 CTCTGAGAGATGACCGGCACTCCAGCTGTCTCTGAGAGAGAGAGAGGAGGAG 1260
DB 1272 CTCTGAGAGATGACCGGCACTCCAGCTGTGTCTCTGAGAGAGAGAGAGGAGGAG 1331
QY 1261 ACAACCAAGCTGAGATCTTAAAGACCATCTCAGAAATCTTCGCGCCCAAGTTCTG 1320
DB 1332 ACAACCAAGCTGAGATCTTAAAGACCATCTCAGAAATCTTCGCGCCCAAGTTCTG 1391
QY 1321 AACCTGTACGAGCTGGAAGGAGAGGCTTCTTGAAGATCTTGTCTCATGACCACTA 1380
DB 1392 AACCTGTACGAGCTGGAAGGAGAGGCTTCTTGAAGATCTTGTCTCATGACCACTA 1451
QY 1381 CTGAGCACCAGAGCCCTCACCAAGAGAGTGTGTCTCTGACAGAGGCTGTGCC 1440
DB 1452 CTGAGCACCAGAGCCCTCACCAAGAGAGTGTGTCTCTGACAGAGGCTGTGCC 1511
QY 1441 AAGGACAGTGGGTGTGAACATGAGACCTGTGTGTGTGAGAGAGATTGAGACGAG 1500
DB 1512 AAGGACAGTGGGTGTGAACATGAGACCTGTGTGTGTGAGAGAGATTGAGACGAG 1571
QY 1501 AACTTTGGGAGATGTTACGCGGACGCTGCGAGCCGACAAACCTGTGTGGGTGAG 1560
DB 1572 AACTTTGGGAGATGTTACGCGGACGCTGCGAGCCGACAAACCTGTGTGGGTGAG 1631
QY 1561 TCTTTGCGAGAGCGCTCCACCTGACCTCAAGGCGCAAGTTTCAAGGAGAGAGATC 1620
DB 1632 TCTTTGCGAGAGCGCTCCACCTGACCTCAAGGCGCAAGTTTCAAGGAGAGATC 1691
QY 1621 CTGAGCAGTACAGCACCACCAATGTGTGTCTGATTTGTGTGCAACCAAGAGCAG 1680
DB 1692 CTGAGCAGTACAGCACCACCAATGTGTGTCTGATTTGTGTGCAACCAAGAGCAG 1751
QY 1681 CCATCTTACATGCTATGAGAGCTTGTGAGAGGAGGAGGAGCTTCTGACCTTCCGACG 1740
DB 1752 CCATCTTACATGCTATGAGAGCTTGTGAGAGGAGGAGGAGCTTCTGACCTTCCGACG 1811
QY 1741 GAGGGGCGCGCTGCGGAGTGAAGCTGTGCTGAGATGATGAGGAGATGAGCTGTGCG 1800
DB 1812 GAGGGGCGCGCTGCGGAGTGAAGCTGTGCTGAGATGATGAGGAGATGAGCTGTGCG 1871
QY 1801 ATGAGTACCTGAGAGCAAGTGTGATCAACCGGAGACCTGTGCTGTGAACTGTCTG 1860
DB 1872 ATGAGTACCTGAGAGCAAGTGTGATCAACCGGAGACCTGTGCTGTGAACTGTCTG 1931
QY 1861 GTGACAGAGAGAGATCTCTGAAGATCAGTGAATCTTGGGATGTCCGAGAGAGAGCGAT 1920

Db	1932	GTGACAGAGGAAGAAATGTCCTGAAATCAGTACCTTTGGGATGTCCCGAGAGGAAGCCGAT	1991
QY	1921	GGGGTCTATGTGACGCTCAGGGGGGCTCAGACAATGTCCTCCCGTGAAGTGAACGGACCTGAG	1980
Db	1992	GGGGTCTATGTGACGCTCAGGGGGGCTCAGACAATGTCCTCCCGTGAAGTGAACGGACCTGAG	2051
QY	1981	GCCCTTAACTACGGGCGGCTACTCCCTCCGAAAGCAGTGTGAGCTTTGGCATCTTGCTC	2048
Db	2052	GCCCTTAACTACGGGCGGCTACTCTCCGAAAGCAGTGTGAGCTTTGGCATCTTGCTC	2111
QY	2041	TGGGAGACTTTCAGCTGGGGGCTCCCTCATATCCCAACTTCAGCATCAGACAACGAG	2100
Db	2112	TGGGAGACTTTCAGCTGGGGGCTCCCTCATATCCCAACTTCAGCATCAGACAACGAG	2171
QY	2101	GAGTTTGTGAGGAAGGGGGGCGCTGTGCTCGCCGACGAGCTGTGCTCGATGCGCGTCTC	2169
Db	2172	GAGTTTGTGAGGAAGGGGGGCGCTGTGCTCGCCGACGAGCTGTGCTCGATGCGCGTCTC	2231
QY	2161	AGGCTCATGTGAGCAGTGTGCGCCTATGAGCCTGGGCAAGCGGCCCAAGCTTCAGACCATC	2220
Db	2232	AGGCTCATGTGAGCAGTGTGCGCCTATGAGCCTGGGCAAGCGGCCCAAGCTTCAGACCATC	2291
QY	2221	TACCAAGAGCTGCAGAGCATCCGAAAGGGGATCGG	2256
Db	2292	TACCAAGAGCTGCAGAGCATCCGAAAGGGGATCGG	2327

```

RESULT 3
US-10-240-965-256
: Sequence 256, Application US/10240965
: Publication No. US20030165924A1
: GENERAL INFORMATION:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: SHIEPMAN, Dov
: APPLICANT: SOMOGYI, Roland
: APPLICANT: LAWN, Richard M.
: APPLICANT: SEILHAMER, Jeffrey J.
: APPLICANT: PORTER, Gordon J.
: APPLICANT: MIMITA, Thomas
: APPLICANT: TAI, Julie
: TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
: FILE REFERENCE: PA-0025 PCT
: CURRENT APPLICATION NUMBER: US/10/240,965
: CURRENT FILING DATE: 2002-10-04
: PRIOR APPLICATION NUMBER: 60/195,106
: PRIOR FILING DATE: 2000-04-05
: NUMBER OF SEQ ID NOS: 276
: SOFTWARE: PERL Program
: SEQ ID NO 256
: LENGTH: 2889
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No. US20030165924A1 997347.6
US-10-240-965-256

```

Query Match	89.7%	Score 2023.4;	DB 16;	Length 2889;
Best Local Similarity	91.4%;	Pred. No. 0;		
Matches 2255; Conservative	0;	Mismatches 1;	Indels 211;	Gaps 2;

Qy	1	ATGGGCTTCTCTTCGAGCTGACGCCCCAGGGCCACGGGGTCTGACGAATGAG	60
Db	188	ATGGGCTTCTCTTCGAGCTGACGCCCCAGGGCCACGGGGTCTGACGAATGAG	247
Qy	61	GAGGCCGAGCTTCGTCATCGGAGGGCATGAGAAATGGATGAGCCACGGGTCAAGAGT	120
Db	248	GAGGCCGAGCTTCGTCATCGGAGGGCATGAGAAATGGATGAGCCACGGGTCAAGAGT	307
Qy	121	GACAGGGAGTATGCAGGACCTGCTTACCACAATATGCCGACAGGACAGTGGGGCCAGAGC	180
Db	308	GACAGGGAGTATGCAGGACCTGCTTACCACAATATGCCGACAGGACAGTGGGGCCAGAGC	367

Oy	181	CGGGCCATTACAGCCCTTGACACACCCCATATGATCAGCTCGGGCTGAGATCAACAAGCCAACT	240
Db	368	CGGGCCATTACAGCCCTTGACACCCCATATGATCAGCTCGGGCTGAGATCAACAAGCCAACT	427
Oy	241	GA -GGGCCTGAGCCGCTTGTCTGCGGACGACGCAAGAGATCTGAATCTAGAGGCCCTTGAG	299
Db	428	GAGGGGCTTAGAGCCGCTGTGCTGCGGACGACGCAAGAGATCTGAATCTAGAGGCCCTTGAG	487
Oy	300	CAAGCTGAGGCTGCTGATCCGGGAAAGGGACGAGCTTCGGAAGACCTACAGCGAGCAGTGTG	359
Db	488	CAAGCTGAGGCTGCTGATCCGGGAAAGGGACGAGCTTCGGAAGACCTACAGCGAGCAGTGTG	547
Oy	360	GCAAGCAGCTGCAAGAGAGCTCAACAAAGCCCAAGCCAGACGACATTTAGAGAGCTGAAGAG	419
Db	548	GCAAGCAGCTGCAAGAGAGAGCTCAACAAAGCCCAAGCCAGACGACATTTAGAGAGCTGAAGAG	607
Oy	420	CCAGTACCCGAGCTCTGGCAACGGGACAGTGGCCCAAGCCCAAGCCAGATCCAGAGAGGCGAG	479
Db	608	CCAGTACCCGAGCTCTGGCAACGGGACAGTGGCCCAAGCCCAAGCCAGATCCAGAGAGGCGAG	667
Oy	480	CAAGAACAAGAACCGGTACAAAGGCAAGGCAAGATAGTGTGCGACCTGTGTGAAGCTCTT	539
Db	668	CAAGAACAAGAACCGGTACAAAGGCAAGGCAAGATAGTGTGCGACCTGTGTGAAGCTCTT	727
Oy	540	TGCTCACACAAACCGCTATGTGTGTGCGGAGCTGCGAGCTACACCAACGACACCA	599
Db	728	TGCTCACACAAACCGCTATGTGTGTGCGGAGCTGCGAGCTACACCAACGACACCA	787
Oy	600	CCACCAAGCTCTGTGCGCGGGCCGTGCGGGGTACATGCAAGAACCTTGACAGAGAGATAGGC	659
Db	788	CCACCAAGCTCTGTGCGCGGGCCGTGCGGGGTACATGCAAGAACCTTGACAGAGAGATAGGC	847
Oy	660	TTGCATCTGTAAGAGAGATCTCTGCAGGAATACCTGAGATTAGCAGCCTGTGTGCAAGATGA	719
Db	848	TTGCATCTGTAAGAGAGATCTCTGCAGGAATACCTGAGATTAGCAGCCTGTGTGCAAGATGA	907
Oy	720	GGTGTGTGGCCATTCAACCGGAGATGTGTGTGCAAGCTGTGCCCCGATCCAGCTTAGAGCTGA	779
Db	908	GGTGTGTGGCCATTCAACCGGAGATGTGTGTGCAAGCTGTGCCCCGATCCAGCTTAGAGCTGA	967
Oy	780	GTACCAAGGCTTCTCTGCGACAGTATGGGTCCGCACTGACGATCCCAACCTGTGTACAGTT	839
Db	968	GTACCAAGGCTTCTCTGCGACAGTATGGGTCCGCACTGACGATCCCAACCTGTGTACAGTT	1027
Oy	840	CGATGATCTCACTGCTTGAGAGGGGTGAACCGCTGGAAGCTTGAGGAGCTCCAGCTGAACGA	899
Db	1028	CGATGATCTCACTGCTTGAGAGGGGTGAACCGCTGGAAGCTTGAGGAGCTCCAGCTGAACGA	1087
Oy	900	GCTGATCTGTGAGAGGCGTGAAGCAACAGCTGAACCTCAAGTACAGATAGAGCTGTGTGTGGC	959
Db	1088	GCTGATCTGTGAGAGGCGTGAAGCAACAGCTGAACCTCAAGTACAGATAGAGCTGTGTGTGGC	1147
Oy	960	CACCGAGATGGTGTTCAGGCGGACAGAGATGGTTACGACGCTGCAACAGAGACTCCGAAA	1019
Db	1148	CACCGAGATGGTGTTCAGGCGGACAGAGATGGTTACGACGCTGCAACAGAGACTCCGAAA	1207
Oy	1020	TGAAGAGAGAAACACCCACCCCCCGGAGACGGGTGCAAGCTGTGGGCAAGAGGCAAGTGTCT	1079
Db	1208	TGAAGAGAGAAACACCCACCCCCCGGAGACGGGTGCAAGCTGTGGGCAAGAGGCAAGTGTCT	1267
Oy	1080	GCAAGAAAGCACTGAGAGGGGCTGACAGTATGAGGCTGTGACGCAAGGCAAGAGCTGACAGGCCA	1139
Db	1268	GCAAGAAAGCACTGAGAGGGGCTGACAGTATGAGGCTGTGACGCAAGGCAAGAGCTGACAGGCCA	1327
Oy	1140	GCAAGAGTTGTGTGAGACCAAGCTGAGACACTTGTGGGCCCGGCGAGGCCCGGCTGTGTCT	1199
Db	1328	GCAAGAGTTGTGTGAGACCAAGCTGAGACACTTGTGGGCCCGGCGAGGCCCGGCTGTGTCT	1387
Oy	1200	GCTCTCTGCAAGATATACCGCCCACTTCCAGTGTCTCTCGGACCAAGAGGAGAGGGGGGAG	1255
Db	1388	GCTCTCTGCAAGATATACCGCCCACTTCCAGTGTCTCTCGGACCAAGAGGAGAGGGGGGAG	1447

Accession	Sequence	Position
Dd	GCAGAAAGCACTGAGGGGGCTGCAGGTAGGGCTGTGCAGCCAGGCCCAAAGCTGCAGGCCCA	1327
Qy	1140 GCAGAGTTGCTGCAGACCAAGCTGAGCACTTGGGCCCGCCGACGCCGCCGCTGTGCT	1199
Dd	1328 GCAGAGTTGCTGCAGACCAAGCTGAGACCACTTGGGCCCGCCGACGCCGCCGCTGTGCT	1387
Qy	1200 GCTCTCTCAGGATGACCGCACTTCCACGTGCTCTTCGAGCAGGAGCGAGAGGGGGGAAG	1259
Dd	1388 GCTCTCTCAGGATGACCGCACTTCCACGTGCTCTTCGAGCAGGAGCGAGAGGGGGGAAG	1447

QY 1260 GAGACCCAGCTGAGATCCTTAAGAGCCATCTCAGAAATCTCCGCCCAAGTTCTC 1319
DB 1448 GAGACCCAGCTGAGATCCTTAAGAGCCATCTCAGAAATCTCCGCCCAAGTTCTC 1507
QY 1320 G----- 1320
DB 1508 GCTCCCTCCACCGCTGAGCTCATTCGAGAGTGACAGAACCCCTGATGAGAGCTGTG 1567
QY 1321 ----- 1320
DB 1568 GTACCAAGGAGCCATCCGAGGAGAGAGTGGCTGAGCTGTGATGACTGTGGAGACTT 1627
QY 1321 ----- 1320
DB 1628 CCTGTGCGGAGAGCCAGGGCAGAGAGTACGTGCTGTGCTGTGCTGTGCTGTGCT 1687
QY 1321 -----AACCTGTACCGACTGAGAGAGAGAGCTT 1349
DB 1688 GCGCCGCGCACTTCATCATCCAGTCTTGATTAACCTGTACCGACTGAGAGAGAGCTT 1747
QY 1350 TCCTAGCATCTTTGCTCATGACCACTTACGAGCAGCAGCAGCAGCAGCAGCAGCAG 1409
DB 1748 TCCTAGCATCTTTGCTCATGACCACTTACGAGCAGCAGCAGCAGCAGCAGCAGCAG 1807
QY 1410 GAGTGTGTGTCTGACAGAGGCTGTGCGCAGAGCAAGTGGGTGTGTAACATGAGGA 1469
DB 1808 GAGTGTGTGTCTGACAGAGGCTGTGCGCAGAGCAAGTGGGTGTGTAACATGAGGA 1867
QY 1470 CTTGTGTGTGTGACAGATTGAGAGGAGGAACTTTGAGGAACTTACGAGCAGCCT 1529
DB 1868 CTTGTGTGTGTGACAGATTGAGAGGAGGAACTTTGAGGAACTTACGAGCAGCCT 1927
QY 1530 GCGAGCCGACACACCTGTGTGCGGTAAGTCTTGTGAGAGACGCTCCACCTGACT 1589
DB 1928 GCGAGCCGACACACCTGTGTGCGGTAAGTCTTGTGAGAGACGCTCCACCTGACT 1987
QY 1590 CAAGGCGAAGTTCTACAGAAAGAGAGTCTGAGAGTCTGAGAGAGAGTCTGAGAG 1649
DB 1988 CAAGGCGAAGTTCTACAGAAAGAGAGTCTGAGAGTCTGAGAGAGAGTCTGAGAG 2047
QY 1650 GCGCTCATTTGTGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1709
DB 2048 GCGCTCATTTGTGTGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2107
QY 1710 GGGGGGCGACTTCTGACTTCTCTCCGACGAGAGGGGGCCGCTGCGGGTGAAGCTCT 1769
DB 2108 GGGGGGCGACTTCTGACTTCTCTCCGACGAGAGGGGGCCGCTGCGGGTGAAGCTCT 2167
QY 1770 GCTGAGATGTGTGGGAGATGACAGTGTGAGATGAGTACTGAGAGAGAGAGTGTGAT 1829
DB 2168 GCTGAGATGTGTGGGAGATGACAGTGTGAGATGAGTACTGAGAGAGAGAGTGTGAT 2227
QY 1830 CCACCGGAGAGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1889
DB 2228 CCACCGGAGAGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2287
QY 1890 TGACTTTGGGATGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1949
DB 2288 TGACTTTGGGATGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2347
QY 1950 ACAAGTCTCCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2009
DB 2348 ACAAGTCTCCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2407
QY 2010 AACGAGAGTGTGAGCTTTGGGATCTTGTCTGTGAGAGAGAGAGAGAGAGAGAGAG 2069
DB 2408 AACGAGAGTGTGAGCTTTGGGATCTTGTCTGTGAGAGAGAGAGAGAGAGAGAGAG 2467
QY 2070 CTATCCCAACTGAGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2129
DB 2468 CTATCCCAACTGAGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2527
QY 2130 CTGCGCAGAGCTGTGTCTGTGATGCGGTGTTGAGAGCTATGAGAGAGAGTGTGCGCCTATGAG 2189

DB 2528 CTGCGCAGAGCTGTGTCTGTGATGCGGTGTTGAGAGCTATGAGAGAGAGTGTGCGCCTATGAG 2587
QY 2190 GCGTGGGAGAGCGGCGGAGCTTACAGACCAATCTTACAGAGAGAGAGAGAGAGAGAGAG 2249
DB 2588 GCGTGGGAGAGCGGCGGAGCTTACAGACCAATCTTACAGAGAGAGAGAGAGAGAGAGAG 2647
QY 2250 GCATCGG 2256
DB 2648 GCATCGG 2654

RESULT 4

US-09-918-995-1503
; Sequence 1503, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918, 995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235, 076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1503
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(449)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-1503

Query Match

Best Local Similarity 99.4%; Pred. No. 3, 9e-72; Length 449;
Matches 328; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1707 GCGAGGGGGGCGACTTCTGACTTCTCTCCGACGAGAGGGGGCCGCTGCGGGTGAAGAG 1766
DB 120 GCGAGGGGGGCGACTTCTGACTTCTCTCCGACGAGAGGGGGCCGCTGCGGGTGAAGAG 179
QY 1767 TCTGCTGAGATGTGTGGGAGATGACAGTGTGAGATGAGTACTGAGAGAGAGTGTGAT 1826
DB 180 TCTGCTGAGATGTGTGGGAGATGACAGTGTGAGATGAGTACTGAGAGAGAGTGTGAT 239
QY 1827 CATCCACCGGAGAGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1886
DB 240 CATCCACCGGAGAGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 239
QY 1887 CAGTGACTTTGGGATGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1946
DB 300 CAGTGACTTTGGGATGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
QY 1947 CAGACAGTCTCCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2006
DB 360 CAGACAGTCTCCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 419
QY 2007 CGAAGCGAGAGTGTGAGCTTTGGGATCTT 2036
DB 420 CGAAGCGAGAGTGTGAGCTTTGGGATCTT 449

RESULT 5

US-09-948-802-5
; Sequence 5, Application US/09948802
; Publication No. US20020172981A1
; GENERAL INFORMATION:
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND

TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
FILE REFERENCE: NMI-090
CURRENT APPLICATION NUMBER: US/09/948,802
CURRENT FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 09/387,212
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 361
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-948-802-5

Query Match 13.1%; Score 296.2; DB 9; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-64;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1613 CGAGATCTCTGAAGAGTACAGCCACCCCAACATCGTCTCATTTGTCGTGACCC 1672
DB 20 CNAAGATCTCTGAAGAGTACAGCCACCCCAACATCGTCTCATTTGTCGTGACCC 79
QY 1673 AGAAGACGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGCACTTCTGACCTTCC 1732
DB 80 AGAAGACGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGCACTTCTGACCTTCC 139
QY 1733 TCCGACCGAGAGGGGGCCCTGCGGGGTGAAGACTCTGTCAGATGTGGGGATGACAG 1792
DB 140 TCCGACCGAGAGGGGGCCCTGCGGGGTGAAGACTCTGTCAGATGTGGGGATGACAG 199
QY 1793 CTCTGGCATGAGTACTCTGAGAGCAAGTCTGACATCCACGGGACCTGCTGCTCGA 1852
DB 200 CTCTGGCATGAGTACTCTGAGAGCAAGTCTGACATCCACGGGACCTGCTGCTCGA 259
QY 1853 ACTGCTGTGTGACAGAGAAATGTCTCTGAAG-ATCAGTACTTT-GGATGTCCGAGA 1910
DB 260 ACTGCTGTGTGACAGAGAAATGTCTCTGAAGATCAGTACTTTGGGATGTCCGAGA 319
QY 1911 GGAAGCC--GATGGGTCTATGACGCTTCAAGGGGGCTTCAAG 1949
DB 320 GGAAGCCGATTTGGGGTCTATGACGCTTCAAGGGGGCTTCAAG 361

RESULT 6
US-10-121-925-5
Sequence 5, Application US/10121925
Publication No. US20030104505A1
GENERAL INFORMATION:
APPLICANT: ROBINSON, KEITH E.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
FILE REFERENCE: NMI-090
CURRENT APPLICATION NUMBER: US/10/121,925
CURRENT FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: US/09/948,802
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: 09/387,212
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 361
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-10-121-925-5

Query Match 13.1%; Score 296.2; DB 15; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-64;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1613 CGAGATCTCTGAAGAGTACAGCCACCCCAACATCGTCTCATTTGTCGTGACCC 1672
DB 20 CNAAGATCTCTGAAGAGTACAGCCACCCCAACATCGTCTCATTTGTCGTGACCC 79
QY 1673 AGAAGACGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGCACTTCTGACCTTCC 1732
DB 80 AGAAGACGCCCATCTACATGTCATGAGCTTGTGACAGGGGGGCACTTCTGACCTTCC 139
QY 1733 TCCGACCGAGAGGGGGCCCTGCGGGGTGAAGACTCTGTCAGATGTGGGGATGACAG 1792
DB 140 TCCGACCGAGAGGGGGCCCTGCGGGGTGAAGACTCTGTCAGATGTGGGGATGACAG 199
QY 1793 CTCTGGCATGAGTACTCTGAGAGCAAGTCTGACATCCACGGGACCTGCTGCTCGA 1852
DB 200 CTCTGGCATGAGTACTCTGAGAGCAAGTCTGACATCCACGGGACCTGCTGCTCGA 259
QY 1853 ACTGCTGTGTGACAGAGAAATGTCTCTGAAG-ATCAGTACTTT-GGATGTCCGAGA 1910
DB 260 ACTGCTGTGTGACAGAGAAATGTCTCTGAAGATCAGTACTTTGGGATGTCCGAGA 319
QY 1911 GGAAGCC--GATGGGTCTATGACGCTTCAAGGGGGCTTCAAG 1949
DB 320 GGAAGCCGATTTGGGGTCTATGACGCTTCAAGGGGGCTTCAAG 361

RESULT 7
US-10-280-576-19
Sequence 19, Application US/10280576
Publication No. US20040044405A1
GENERAL INFORMATION:
APPLICANT: WOLFE, Matthew R.
TITLE OF INVENTION: VASCULAR STENT OR GRAFT COATED OR IMPREGNATED WITH PROTEIN
FILE REFERENCE: 09820.189
CURRENT APPLICATION NUMBER: US/10/280,576
CURRENT FILING DATE: 2002-10-22
PRIOR APPLICATION NUMBER: 60/343,732
PRIOR FILING DATE: 2001-10-25
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 19
LENGTH: 1779
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-10-280-576-19

Query Match 12.0%; Score 271.6; DB 17; Length 1779;
Best Local Similarity 57.0%; Pred. No. 3.6e-58;
Matches 518; Conservative 0; Mismatches 384; Indels 6; Gaps 1;

QY 1327 TACCGATGGAAGGGGAGGCTTCTTAGCATCTTGTCTCATGACCACTAGTACG 1386
DB 844 TTCCGTTGAGGGAGCACCACTTTGCGAGCATCAGAGCTGATCATGATCAGTATCAC 903
QY 1387 ACCCAGACCCCTCTCAACAAGAGTGTGTCTCTGACAGGGCTGTGCCAAGAC 1446
DB 904 TCCGATTTGAGAGTACCGTGAATCGGAGGCACTCTCGACGCCGTTTCCGAG 963
QY 1447 AACTGGGTGTGACCATGAGACTGTGTGTGGGTGAGCAGATTTGAGGGGAACTTT 1506
DB 964 CGCTGGAGCTGTGACCAAGATGATGTGTGATCTTGTGAAGAGATGTGTGGGAACTTT 1023
QY 1507 GCGAAGTGTTCAGCGGAGCGCTTGCAGAGCGACCAACCTGTGTGCGGTGAAGTCTTGT 1566
DB 1024 GGGGATGTCTTACAGGCCAACTGAAGTCCACCACTGATGTGTCTTCAAACTGT 1083
QY 1567 CGAGAGCGTCTCCACCTGACCTTCAAGGCCAATTTCTACAGGAAGCAGAGTCTGAAG 1626
DB 1084 CGAATACCTGTCCGACGAGAGAGGCTTAATTTCTTACAGGAAGGGGCACTCTCAAG 1143
QY 1627 CAGTACGACGCCCAACATGTGTCTGTCTGATGTGTGTGACCCAGAGAGGCCCATC 1686
DB 1144 CAATACGATCATCAATATCTGTAAATGATTTGTGATTTGTGTGAGAGAGGCCCATC 1203

```

Oy      1667  TACATGCGCATGAGAGCTTGTGTGACAGGGGGGACGACTTCTCGGACCTTCTCCGACGAGAGGG 1746
Db      1204  ATGATTTGTCATGGAATTTGTCTCTGGTGTTCCTTTTACTTTATTAACGAAAGACTCC 1263
Oy      1747  GCCCGCCTGCGGGGTGAAGACTCTGCTGCGATGTGTGGGGAGTGCAGCTGCTGCATGAG 1806
Db      1264  AATGGCCCTCACCACTCGGCAACAATGGGCAATGTGCAGATGTGGCGGGCAGCATGGGA 1323
Oy      1807  TACCTTGAGAGCAAGTGTGCATCACCGGGAACTGTGTCTGCGAACTGCTGTGTACA 1866
Db      1324  TATGTGGAGTCCAAAAMATGCAATTCATCGCATCTGGGCGGCGTAATTTGTTGTGAC 1383
Oy      1867  GAGAAAGATGTCCTGAGATCAGTGACTTTTGGATGTCCGAGAGAAAGCCGATGGGGTC 1926
Db      1384  TTGGAGCACAGTGTGAAGATTTCCGATTTCCGAAATGTCTCCGAGAA-----GAGGA 1437
Oy      1927  TATGACGCTGAGGGGGGCTTCAGACAGACTCCCGTGAAGTGAACCGCACTGAGGCCCTT 1986
Db      1438  TATATAGTTTCCGATGGCATGAACAATACTGTGAAGTGAACGCTCCGAGGCCCTTG 1497
Oy      1987  AACACGCGCCCTACTCTCTCCGAAAGGACGTGTGAGCTTTTGGCATTTTGTCTCTGGAG 2046
Db      1498  AATTTGGCAAGTACACTTCTGTGTGCATGTGTGTCTATGTGCAATACGTATGTGGAG 1557
Oy      2047  ACCTTACGCTGTGGGGGCTCCGCCCTATGCCAACTCCAGCAATCAGACAGACAGGAGATT 2106
Db      1558  ATCTTCTTCAAGGGCGACACACCTTACTCTCGGATGACCACTCAGAGCCAGAGAGGCC 1617
Oy      2107  GTGAGAAAGGGGGGCGCTGTGCTGCCCTGCCCAAGAGCTGTGTCTGTATGCCGTTCAGGCTC 2166
Db      1618  ATCATTAACGGGATATCTGTATGCCAAACGCCGAAGAGACAGCCCGAGAGATGTACCGACTG 1677
Oy      2167  ATGAGAGCAGTGTGGGCGCTATGAGCGTGGGAGAGGGCGCCACTTCAGACCATCTACAG 2226
Db      1678  ATGCTCCAGTCTGTGGCAGCCGAGCGCAATCCGACCTTCGATGAGATCTTACAT 1737
Oy      2227  GAGCTGCA 2234
Db      1738  GTGGTGA 1745

RESULT 8
US-10-003-295-3
; Sequence 3, Application US/10003295
; Publication No. US20020168741A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01183D1V
; CURRENT APPLICATION NUMBER: US/10/003,295
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-3

```

Oy	121	GACAGGGAGATATGACGAGATGCTTCA	CCACATGTCCTGACAGGACAGTGGGGGCCAAGC	180
Db	2683	GACAGGGAGATATGACGAGATGCTTCA	CCACATGTCCTGACAGGACAGTGGGGGCCAAGC	2742
Oy	181	CGGGCCATACAGCCCTGACAGCCCATCAGTC	-----	211
Db	2743	CGGGCCATACAGCCCTGACAGCCCATCAGTC	AGTGGAGTCTGTGGGT	2802
Oy	212	-----	-----	211
Db	2803	GCTGGCGTATCTGCTTCTCTCTCTCTCT	GCGGGGCCCTCTGGGGGACATGAGCTGAGAT	2862
Oy	212	-----	-----	211
Db	2863	CTGGCAGGGCCATGCTTGGGAGCCATGTG	CCCCCTCCCTGCTCCCATCTGTCTG	2922
Oy	212	---AGTCTTGGGCTGAGATCA	CCAGCCAACTGAGGGGCTGAGCGCGTGTCTGGGGACG	268
Db	2923	TATGTGCTTGGGCTGAGATCA	CCAGCCAACTGAGGGGCTGAGCGCGTGTCTGGGGACG	2982
Oy	269	ACGACAGAGATCTGAATCTCAGGGGCCCT	GACAGTGAAGCTCTGATCCGGGAAAGGC	328
Db	2983	ACGACAGAGATCTGAATCTCAGGGGCCCT	GACAGTGAAGCTCTGATCCGGGAAAGGC	3042
Oy	329	AGCAGCTTCCGACAGACTTACAGCGACG	ATGAGCAGCTTGCACAGAGAGCTTCA	388
Db	3043	AGCAGCTTCCGACAGACTTACAGCGACG	ATGAGCAGCTTGCACAGAGAGCTTCA	3102
Oy	389	CCCAAGCAGAGGAC	402	
Db	3103	TGAGCGGGCAGCAC	3116	

```

RESULT 9
US-10-660-763-3
; Sequence 3, Application US/10660763
; Publication No. US20040063130A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C000183DIV1
; CURRENT APPLICATION NUMBER: US/10/660.763
; CURRENT FILING DATE: 2003-09-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-660-763-3

```

```

: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 15297
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-003-295-3

Query Match      10.1% Score 227.2; DB 13; Length 15297;
Beet Local Similarity 71.1%; Pred. No. 7.3e-47;
Matches 394; Conservative 0; Mismatches 8; Indels 152; Gaps 1;

Qy      1  ATGGGCTTCCTTCTGAGCTGTGACCCCGAGGCGCCAGCGGAGTCTCTCAGCAAAATGCAG 60
Db      2563  ATGGGCTTCCTTCTGAGCTGTGACCCCGAGGCGCCAGCGGAGTCTCTCAGCAAAATGCAG 26222

Qy      61  GAGGCGCGAGCTTCCTGCTCTACTGAGGAGGCGCATGAGAAAGTGAGTGGCCCAAGGGGTCAAGAGT 120
Db      2623  GAGGCGCGAGCTTCCTGCTCTACTGAGGAGGCGCATGAGAAAGTGAGTGGCCCAAGGGGTCAAGAGT 26822

```

Query Match	Similarity	10.1%	Score 227.2	DB 17	Length 15297
Best Local	Similarity	71.1%	Pred. No. 7.3e-47		
Matches 394	Conservative	0	Mismatches 8	Indels 152	Gaps 1
QY	1	ATGGGCTTCTTCTTGAGCTGTGACGCCCCGAGGCCACGGAGTCTCTGACGAATATGAG	60		
Db	2563	ATGGGCTTCTTCTTGAGCTGTGACGCCCCGAGGCCACGGAGTCTCTGACGAATATGAG	26222		
QY	61	GAGGCCGAGCTTCTCTATCTGAGAGGGCATGAGAAAGTGAATGGGCCGAGCGGGGTCAAGAT	120		
Db	2623	GAGGCCGAGCTTCTCTATCTGAGAGGGCATGAGAAAGTGAATGGGCCGAGCGGGGTCAAGAT	26822		
QY	121	GACAGGGAGTATGACAGGACTGCTTACCAACAATGTCTCTTGACAGGACAGTGGGGCCACAGC	180		
Db	2683	GACAGGGAGTATGACAGGACTGCTTACCAACAATGTCTCTTGACAGGACAGTGGGGCCACAGC	27422		
QY	181	CGGGCCATCAACCTCTGACAGCCCATCATGTC-----	211		
Db	2743	CGGGCCATCAACCTCTGACAGCCCATCATCATGATGGTGGTCTCTATAGGACTCTGTGGGT	28022		

QY 212 ----- 211
Db 2803 GCTGGGATCTGCTTCTCTCTCTCTGGGGGCTCTGGGGGAGTGGCTGGAGAT 2862
QY 212 ----- 211
Db 2863 CTGGAGAGCCAAATGTTGGAGGCAATGTGTGCCCCCTCTGCTGCTCCCATCTGTGTG 2922
QY 212 ---AGTCTGGGCTGAGATACCAAGCCAACTGAAGGCTGAGCCGCTTGTCTGGGAGC 2868
Db 2923 TATATGCTGGGCTGAGATACCAAGCCAACTGAAGGCTGAGCCGCTTGTCTGGGAGC 2982
QY 269 AGCAGAGAGATCTGAATCTGAGGCCCCCTGAGCAAGCTGAGCTGATCTGATCCGGAACGGC 328
Db 2983 AGCAGAGAGATCTGAATCTGAGGCCCCCTGAGCAAGCTGAGCTGATCTGATCCGGAACGGC 3042
QY 329 AGCAGCTTGGCAAGCCTTACAGGAGAGAGTGGAGAGCTGAGCAAGAGCTTACCAAGA 388
Db 3043 AGCAGCTTGGCAAGCCTTACAGGAGAGAGTGGAGAGCTGAGCAAGAGCTTACCAAGG 3102
QY 389 CCCACAGCCGAGAC 402
Db 3103 TGAGCGGCGAGCAGC 3116

RESULT 10

US-10-101-510-525
; Sequence 525, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: MAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 525
; LENGTH: 3875
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-525

Query Match 8.9%; Score 201.2; DB 15; Length 3875;
Best Local Similarity 54.9%; Pred. No. 2e-40;
Matches 423; Conservative 0; Mismatches 338; Indels 9; Gaps 1;

QY 1451 GGGTGTGGAACCATGAGAGACCTGTGTGGGTGAGCAGATTTGACGGGGAACTTTGGCG 1510
Db 1969 GGGAGCTTGTATCCAGCTGTGCTGATGTGAGACCTGTCTATGAGAAAGAGAGATTGGGG 2028
QY 1511 AAGTGTTCAGCGGAGCGCTCGAGCC-----GACAAACCTTGTGGCGGTGAAGT 1561
Db 2029 AAGTGTATGAGGAGACCTTGAAGCTCCCGACGAGACTGCTGAGCTTGTGGAG 2088
QY 1562 CTGTGCGAGAGCGCTTCCACTGACTCAAGGCCAAATTCTTACAGAAAGAGAGATCC 1621
Db 2089 CTTTAAAGACATCTCCAGTGGCCAGTGTGAACTTCTTTCAGAGGCAATATCA 2148
QY 1622 TGAAGCAGTACAGCCACCCCAACATGCTGTCTCTATGCTGTCTGACCCAGAGAGC 1681
Db 2149 TGGGCCAGTTTACCCACCCCATATTTCTGATCTGGAAGGCGTGTCAAAACGAAAGC 2208
QY 1682 CCATTCATCATGCTGATGAGCTTGTGCAAGGAGGCGCACTTCTGAGCTTCTCGGACGG 1741
Db 2209 CGATTCATCATCATCAAGAAATTTATGGAATGAGAGCCCTTGTGATGCTTCTGAGGAGC 2268
QY 1742 AGGAGGCGCGCTGCGGGTGAAGACTTGTCTGAGATGTGGGGAGATGAGCTGTGGCA 1801

Db 2269 GGGAGAGCAGCTGTGCTCTGGGCACTAGTGCCATGCTGAGGAGGATAGCATCTGGCA 2328
QY 1802 TGAGTACCTGAGAGCAAGTCTGCATCCACGGGAGCTTGGCTGTGCGAACTGCTGG 1861
Db 2329 TGAATCTCTGAGTATCAAAATTAATGTCCACGGGAGCTTGGCTGCCAAGAACTTTGG 2388
QY 1862 TGACAGAGAAAGATGCTGAAAGATCAGTACTTGTGGATGTGCCAGAGAGAGCCGATG 1921
Db 2389 TGAATCAAAACCTGTGCTCAAGGTGTGTGACTTTGGCTGTGACTGCTCTCTGATGACT 2448
QY 1922 GGGTCTATGACCTTCAGGGGGGCTTCAGCAAGTCCCGCTGAAGTGAACCCGACTGAG 1981
Db 2449 TTGATGGCAATATCAAGAACCCAGGAGAGAAATCTCTATCCGTGGAGAGCCCTGAAG 2508
QY 1982 CCTTAACTAGCGCGCTACTCTCCGAAAGAGAGCTTGGAGCTTGGCATCTTGTCT 2041
Db 2509 CATTTGCCATGAGATCTTACACAGCCAGGAGATGTGAGCTTTGGGATTTGTATGT 2568
QY 2042 GGGAGACTTTCAGCTGAGGGGCTCCCTTATCCAACTCAGCAATCAGACAGCGG 2101
Db 2569 GGGAGGTGCTGAGCTTTGGGGAAGAGCTTATGGGAGATGAGCAATCAGAGATTATGA 2628
QY 2102 AGTTTGTGAAGAAAGGGGGGCTGTGCTGCTGCCAGAGCTGTGTCTGATGCCGTGTA 2161
Db 2629 AGAGCATGAGAGATGGGTACCGGTTGCCCTCTGTGTGAGCTGCTGCGCCCTGTGTATG 2688
QY 2162 GGGTCATGAGAGAGTGTGGGCTTATGAGCTTGGGAGGCGCCAGCTTC 2211
Db 2689 AGCTCATGAAGAACTGTGGGCAATGACCGTGTGCCCGGCGGCAACTTC 2738

RESULT 11

US-10-384-339C-1
; Sequence 1, Application US/10384339C
; Publication No. US20040175703A1
; GENERAL INFORMATION:
; APPLICANT: Kreutzler, Roland
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITING EXPRESSION OF A TARGET GE
; FILE REFERENCE: 20200/2002
; CURRENT APPLICATION NUMBER: US/10/384,339C
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/EP02/00152
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: DE 10100586.5
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: DE 10155280.7
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: DE 10158411.3
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: DE 10160151.4
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2955
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; TITLE: EPN A1
; PATENT DOCUMENT NUMBER: NM00532
US-10-384-339C-1

Query Match 8.8%; Score 198; DB 18; Length 2955;
Best Local Similarity 54.7%; Pred. No. 1.3e-39;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1451 GGGTGTGGAACCATGAGAGCTGTGTGGGTGAGCAGATTTGACGGGGAACTTTGGCG 1510
Db 1874 GGGAGCTTATCCAGGCTGTGATGTGTGAGACCTGTCTATGAGAAAGAGAGATTGGGG 1933
QY 1511 AAGTGTTCAGCGGAGCGCTCGAGCC-----GACAAACCTTGTGGCGGTGAAGT 1561
Db 1934 AAGTGTATGAGGAGACCTTCAGGCTCCCGACGAGACTGCAAGACTGTGGCAATTAAGA 1993

QY 1562 CTTGTGAGAGAGGCTCCCACTGACCTCAAGGCCAAGTTTCTTACAGAAAGCAGATCC 1621
DB 1994 CTTAAAGACATCCCAAGGAGGCTGAGTGTGAACTTCTTCAAGAGCAATATCA 2053
QY 1622 TGAAGCAGTACAGCCCAACCATGCTGCTCTCAATTGGTGTCTCAACCCAGAAACAG 1681
DB 2054 TGGGCCAGTTTGAACCCGCAATTTCTGATCTGGAAGGCGTGTCAACAAAGCAAG 2113
QY 1682 CCATCTACATCGTCAATGAGCTTGTGCAAGGGGCGAATTCTGACCTTCTCCGACGG 1741
DB 2114 CGATCATGTATCAACAGATTATAGAGATCAACCCCTGATGATGCTTCTGAGGAGC 2173
QY 1742 AGGGGGCCCGCTGCGGGTGAAGACTGTGTGAGATGAGTGGGGATGACAGCTGTGGA 1801
DB 2174 GGAAGGACCAAGCTGTCTCCCTGGGCACTGAGGCAATGCTGCAAGGCAATGACTGGCA 2233
QY 1802 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGAGCTGCTGCTGGAATGCTGCTGG 1861
DB 2234 TGAATCTACCTCAAGTATCAACATTTATGTCCACCGGAGCTGCTGCTGCAAGAAATTTGG 2293
QY 1862 TGAAGAGAAAGATGTCTGGAAGATGATGATCTTTGGATGTCCGAGAGAAAGCCGATG 1921
DB 2294 TGAATCAAAACCTGTGCTGCAAGGTGTCTGACTTTGGCTGCACTGCGCTCTGATGATG 2353
QY 1922 GGGTCTATGAGCTCAAGGGGGGCTCAAGCAAGTCCCGCTGAAGTGAACCGCACTGAGG 1981
DB 2354 TTGATGCACTATGCAAAAGCCAGAGGAAAGATCCCTATCCCTGTGAGACGCCCCCTGAAG 2413
QY 1982 CCCTTAACAGCGCGCTACTCCTCCGAAAGGAGCGTGTGAGCTTTGGATTTGTCTCT 2041
DB 2414 CCATTCCTCATCGGATCTTCAACAGCCAGCAATGTGTGAGACTTTGGATTTGTATGT 2473
QY 2042 GGGAGACTTTCAGCTGCGGGGCTCCCTATCCCACTCAAGCAATCAAGCAGACAGCGG 2101
DB 2474 GGAAGGTGCTGAGCTTTGGGGAACAAGCTTATGAGGAGATGACAAATCAGAGATTATGA 2533
QY 2102 AGTTTGTGAGAAAGGGGGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2161
DB 2534 AGAGCATTTGAGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2593
QY 2162 GGGTCAATGAGAGCTGTGAGCTTATGAGCTGAGGAGCGGCGCCAGCTTC 2211
DB 2594 AGCTCATGAAGAACTGCTGAGGCAATATGACCTGTCCCGCGGCAACTTC 2643

RESULT 12

US-09-967-768A-144

Sequence 144, Application US/09967768A
Patent No. US20020150877A1

GENERAL INFORMATION:

APPLICANT: Augustus, Meena

TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu

FILE REFERENCE: 689290-72

CURRENT APPLICATION NUMBER: US/09/967,768A

CURRENT FILING DATE: 2001-09-28

PRIOR APPLICATION NUMBER: US/60/236,109

PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US/60/236,034

PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US/60/236,111

PRIOR FILING DATE: 2000-09-28

NUMBER OF SEQ ID NOS: 325

SOFTWARE: Patentin version 3.0

SEQ ID NO 144

LENGTH: 3370

TYPE: DNA

ORGANISM: Homo sapiens

US-09-967-768A-144

Query Match 8.8%; Score 198; DB 9; Length 3370;
Best Local Similarity 54.7%; Pred. No. 1.3e-39;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;
QY 1451 GGGTCTGAACCATGAGCACTGTGTGTGGTGTGAGCAATTTGACCGGGAACTTTGGCG 1510
DB 1967 GGGAGCTTATCATCGAGT 2026
QY 1511 AGTGTTCAGGAGGAGCGCTGCGAGCC-----GACAAACCCGTGTGGCGGTGAAGT 1561
DB 2027 AGTGTATGAGGAGACCTTCAGGCTTCCAGCCAGAGATGCAAGACTGTGTGTGTGTGTGT 2086
QY 1562 CTTGTGAGAGAGCGTCCCACTGACCTCAAGGCCAAGTTTCTTACAGAAAGCAGATCC 1621
DB 2087 CTTAAAGACATCCCAAGGAGGCTGAGTGTGAACTTCTTCAAGAGCAATATCA 2146
QY 1622 TGAAGCAGTACAGCCCAACCATGCTGCTCTCAATTGGTGTCTCAACCCAGAAACAG 1681
DB 2147 TGGGCCAGTTTGAACCCGCAATTTCTGATCTGGAAGGCGTGTCAACAAAGCAAG 2206
QY 1682 CCATCTACATCGTCAATGAGCTTGTGCAAGGGGCGAATTCTGACCTTCTCCGACGG 1741
DB 2207 CGATCATGTATCAACAGATTATAGAGATGACAGCCCTGATGATGCTTCTGAGGAGC 2266
QY 1742 AGGGGGCCCGCTGCGGGTGAAGACTGTGCTGAGATGAGTGGGGATGACAGCTGTGGA 1801
DB 2267 GGAAGGACCAAGCTGTCTCCCTGGGCACTGAGGCAATGCTGCAAGGCAATGACTGTGGA 2326
QY 1802 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGAGCTGCTGCTGGAATGCTGCTGG 1861
DB 2327 TGAATCACTCAAGTATCAACATTTATGTCCACCGGAGCTGCTGCTGCAAGAAATCTTGG 2386
QY 1862 TGAAGAGAAAGATGTCTGGAAGATGATGATCTTTGGATGTCCGAGAGAAAGCCGATG 1921
DB 2387 TGAATCAAAACCTGTGCTGCAAGGTGTGACTTTGGCTGCACTGCTCTCTGAGTACT 2446
QY 1922 GGGTCTATGAGCTCAAGGGGCTCAAGCAAGTCCCGCTGAAGTGAACCGCACTGAGG 1981
DB 2447 TTGATGCACTATGCAAAAGCCAGAGGAAAGATCCCTATCCCTGTGAGACGCCCCCTGAAG 2506
QY 1982 CCCTTAACAGCGCGCTACTCCTCCGAAAGGAGCGTGTGAGCTTTGGATTTGTCTCT 2041
DB 2507 CCATTCCTCATCGGATCTTCAACAGCCAGCAAGTGTGTGAGCTTTGGATTTGTATGT 2566
QY 2042 GGGAGACTTTCAGCTGCGGGGCTCCCTATCCCACTCAAGCAATCAAGCAGACAGCGG 2101
DB 2567 GGAAGGTGCTGAGCTTTGGGGAACAAGCTTATGAGGAGATGACAAATCAGAGATTATGA 2626
QY 2102 AGTTTGTGAGAAAGGGGGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2161
DB 2627 AGAGCATTTGAGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2686
QY 2162 GGGTCAATGAGAGCTGTGAGCTTATGAGCTGAGGAGCGGCGCCAGCTTC 2211
DB 2687 AGCTCATGAAGAACTGCTGAGGCAATATGACCTGTCCCGCGGCAACTTC 2736

RESULT 13

US-10-354-358-101

Sequence 101, Application US/10354358

Publication No. US20030157082A1

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc

APPLICANT: Hunter, John Joseph

APPLICANT: MacBeth, Kyle J.

APPLICANT: Tsai, Feng-Ying

APPLICANT: Lesoon, Andrea

APPLICANT: Lightcap, Eric S.

APPLICANT: Williamson, Mark

APPLICANT: Rudolph-Owen, Laura A.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING

TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,

TITLE OF INVENTION: 7181, 7660, 25641, 69583, 8897, 1682, 17667, 9235,

TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,

TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,

TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2093,
TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
FILE REFERENCE: MPI02-020P1RNM01M
CURRENT APPLICATION NUMBER: US/10/354,358
PRIOR FILING DATE: 2003-01-30
PRIOR APPLICATION NUMBER: US 60/353,600
PRIOR FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: US 60/364,517
PRIOR FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/371,075
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: US 60/371,507
PRIOR FILING DATE: 2002-04-10
PRIOR APPLICATION NUMBER: US 60/372,984
PRIOR FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: US 60/374,194
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/382,995
PRIOR FILING DATE: 2002-05-24
PRIOR APPLICATION NUMBER: US 60/385,023
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: US 60/388,853
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/389,395
PRIOR FILING DATE: 2002-06-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 122
SOFTWARE: FaSTSeq for Windows Version 4.0
SEQ ID NO 101
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (94)... (3048)
US-10-354-358-101

Query Match 8.8%; Score 198; DB 16; Length 3370;
Best Local Similarity 54.7%; Pred. No. 1.3e-39;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

1451 GGGTGTGAACATGAGACCTGTGTTGGGTGAGCAGATTGACCGGGGAATTGGCG 1510
1967 GGGAGCTTGATTCAGGCTGTGCTGATGTGACACTGTGAGAGAGAGATTGGGG 2026
1511 AAGTGTTCAGCGGACCGCTCGAGCC-----GACAAACCTGTGTGGCGTGAAGT 1561
2027 AAGTGTATCAGAGGACCTCGAGCTCCCAAGCAGGACTGTCGATTAAGA 2086
1562 CTGTGTGAGAGAGCTCCCACTGAAGCCAAAGTTTCAAGAGAGAGATCC 1621
2087 CCTTAAAGACATCCCGAGGTGGCCAGTGTGGAATCTTCTTCAAGAGGCAATATCA 2146
1622 TGAAGCAGTACAGCAACCCCAACATGCTGCTCATTTGTTCTGACCCAGAACAGC 1681
2147 TGGGCGCAGTTTACGACCGCATATTCATCTGGAAGCGGTGTCAAGAAAGC 2206
1682 CCATCTACATGCTGATGAGCTGTGCAAGGGGCGAATTCCTGACCTTCTCCGACGG 1741
2207 CGATCATGATCATCAAGAAATTTATGAGATCAGCCCTGGAATGCTTCTGAGGAGC 2266
1742 AGGGGGCCGCGCTGCGGGTGAAGACTGTGCTGCAAGATGTTGGGGAATGCACTGTGGCA 1801
2267 GGGAGGACAGCTGTGCTCTGGGCAAGTATGAGCCATGTGCAAGGGCATATGCA 2326
1802 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGGAATCTGTGCTCGAACTGTGG 1861
2327 TGAATCACTCAATATCAATATATGTCCACCGGACCTGTGCTCGAAGAAATCTTGG 2386
1862 TGAAGAGAAATGTCTTGAAGATCAATGACTTTGGGATGTCCGAGAGAAACCGATG 1921
2387 TGAATCAAACTGTGTGCAAGGTGTGTGACTTTGGCTGACTGCGCTCTGTGAATGACT 2446

1922 GGGTGTATCAGACCTCAGGGGCTCAGACAAGTCCCGTGAAGTGAACCGACCTGAG 1981
2447 TGAATGACATACAGAAACCCAGGAGAAATCCCTATCCGTTGACAGCCCTGGAAG 2506
1982 CCTTAACTACGCGCGCTACTCTCCGAAAGCACTGTGAGCTTTGGCATCTTGTCT 2041
2507 CCATGCCCCATGAGATCTTCAACACAGCGAGATGTGAGCTTTGGGATTTGATGT 2566
2042 GGGAGACTTTCAGCCCTGGGGGCTCCCGTATCCCACTCAGCATCAGACAGCGG 2101
2567 GGGAGGTGTGAGCTTTGGGGAAGCTTATGAGGAGATGACAAATCAGAGTTTATTA 2626
2102 AGTTGTGAGAGAGGGGGCGCTGCTGCTGCGCCAGCTGTGTCTGATGCGGTTC 2161
2627 AGAGCATGAGATGAGTATCCGGTTGCCCCCTCTGTGACTGCTGCCCCCTGTATG 2686
2162 GGGTCATGAGAGAGTGTGGGCTTATGAGCTTGGGAGCGGCCACTTC 2211
2687 AGCTCATGAAGAACTGCTGGGATATGACCGTCCCGCCGACACACTTC 2736

RESULT 14
US-10-210-120-19

/ Sequence 19, Application US/10210120
/ Publication No. US20030175736A1

/ GENERAL INFORMATION:

/ APPLICANT: Chinnaiyan, Arul M.

/ APPLICANT: Rubin, Mark A.

/ TITLE OF INVENTION: Expression Profile of Prostate Cancer

/ FILE REFERENCE: UM-07221

/ CURRENT APPLICATION NUMBER: US/10/210,120

/ PRIOR APPLICATION NUMBER: US 60/309,581

/ PRIOR FILING DATE: 2001-08-02

/ PRIOR APPLICATION NUMBER: US 60/334,468

/ PRIOR FILING DATE: 2001-11-15

/ NUMBER OF SEQ ID NOS: 123

/ SOFTWARE: PatentIn version 3.2

/ SEQ ID NO 19

/ LENGTH: 3370

/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-10-210-120-19

Query Match 8.8%; Score 198; DB 16; Length 3370;
Best Local Similarity 54.7%; Pred. No. 1.3e-39;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

1451 GGGTGTGAACATGAGACCTGTGTTGGGTGAGCAGATTGACCGGGGAATTGGCG 1510
1967 GGGAGCTTGATTCAGGCTGTGCTGATGTGACACTGTGAGAGAGAGATTGGGG 2026
1511 AAGTGTTCAGCGGACCGCTCGAGCC-----GACAAACCTGTGTGGCGTGAAGT 1561
2027 AAGTGTATGAGGAGACCTCAGGCTCCCAAGCAGGACTGCAAGATGTTGCAATTAAGA 2086
1562 CTGTGTGAGAGAGCTCCCACTGAAGCCAAAGTTTCAAGAGAGAGATCC 1621
2087 CCTTAAAGACATCCCGAGGTGGCCAGTGTGGAATCTTCTTCAAGAGGCAATATCA 2146
1622 TGAAGCAGTACAGCAACCCCAACATGCTGCTCATTTGTTCTGACCCAGAACAGC 1681
2147 TGGGCGCAGTTTACGACCGCATATTCATCTGGAAGCGGTGTCAAGAAAGC 2206
1682 CCATCTACATGCTGATGAGCTGTGCAAGGGGCGAATTCCTGACCTTCTCCGACGG 1741
2207 CGATCATGATCATCAAGAAATTTATGAGATCAGCCCTGGAATGCTTCTGAGGAGC 2266
1742 AGGGGGCCGCGCTGCGGGTGAAGACTGTGCTGCAAGATGTTGGGGAATGCACTGTGGCA 1801
2267 GGGAGGACAGCTGTGCTCTGGGCAAGTATGAGCCATGTGCAAGGGCATATGCA 2326

QY 1802 TGAAGTACCTGAGAGCAAGTGTGATCCACCGGAACTTGGCTGTGCGAACTGCTCG 1861
DB 2327 TGAACCTGATGATATCAATTAATATGACCGGAACTTGGCTGTGCGAACTGCTCG 2386
QY 1862 TGAACGAGAAAGATGCTGGAAGATGATGATGATGATGATGATGATGATGATGATG 1921
DB 2387 TGAATCAAAACCTGTGCTGCAAGGTGTGATGATGATGATGATGATGATGATGATG 2446
QY 1922 GGGTCTATGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1981
DB 2447 TTGATGAGCACAATGCAAAACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2506
QY 1982 CCCTTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2041
DB 2507 CCATGGCCATGAGGATCTTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2566
QY 2042 GGGAGACCTTCAAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2101
DB 2567 GGGAGGCTGAGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2626
QY 2102 AGTTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2161
DB 2627 AAGAGATGAGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG 2686
QY 2162 GGCTCATGAGAGATGCTGGGCTTATGAGGCTTGGGAGGAGGAGGAGGAGGAGGAG 2211
DB 2687 AGCTCATGAGAGATGCTGGGCTTATGAGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAG 2736

RESULT 15

US-10-776-827-82
; Sequence 82, Application US/10776827
; Publication No. US20040132086A1
; GENERAL INFORMATION:
; APPLICANT: Horwiltz, Jennifer
; APPLICANT: Richer, Jennifer
; TITLE OF INVENTION: Progestosterone Receptor-Regulated Gene Expression and Methods Relat
; FILE REFERENCE: 2848-39
; CURRENT APPLICATION NUMBER: US/10/776,827
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US/09/814,915
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 60/214,870
; PRIOR FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 82
; LENGTH: 3370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-776-827-82

Query Match 8.8%; Score 198; DB 18; Length 3370;

Best Local Similarity 54.7%; Pred. No. 1.3e-39;

Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1451 GGGTGTGTAACCATGAGACCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1510
DB 1967 GGGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2026
QY 1511 AAGTGTTCAGCGGACCTGCGAGCC-----GACACACCTGTGTGCGGTGAGT 1561
DB 2027 AAGTGTATGAGGAGACCTTCAAGGCTCCGACGAGGAGCTGCAAGAGCTGTGGCATTAAGA 2086
QY 1562 CTTGTGAGAGAGCGCTCCACCTGACCTCAAGGCAAGTTTCTACAGAGAGGAGATCC 1621
DB 2087 CCTTAAAGACACATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2146
QY 1622 TGAAGCAGTACAGCAACCCCAACATGATGATGATGATGATGATGATGATGATGATGATG 1681
DB 2147 TGGGCGAGTTTACGCCCGGATATTTCTGATCTGAGAGGCGGTGTCACAAAGCGAAAGC 2206

QY 1682 CCATCTACATGCTCATGAGAGCTTGTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1741
DB 2207 CGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2266
QY 1742 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1801
DB 2267 GGGAGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2326
QY 1802 TGAAGTACCTGAGAGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1861
DB 2327 TGAATCAAAACCTGTGCTGCAAGGTGTGATGATGATGATGATGATGATGATGATGAT 2386
QY 1862 TGAACGAGAAAGATGCTGGAAGATGATGATGATGATGATGATGATGATGATGATG 1921
DB 2387 TGAATCAAAACCTGTGCTGCAAGGTGTGATGATGATGATGATGATGATGATGATGAT 2446
QY 1922 GGGTCTATGACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1981
DB 2447 TTGATGAGCACAATGCAAAACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2506
QY 1982 CCCTTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2041
DB 2507 CCATGGCCATGAGGATCTTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2566
QY 2042 GGGAGACCTTCAAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2101
DB 2567 GGGAGGCTGAGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2626
QY 2102 AGTTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2161
DB 2627 AAGAGATGAGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG 2686
QY 2162 GGCTCATGAGAGATGCTGGGCTTATGAGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAG 2211
DB 2687 AGCTCATGAGAGATGCTGGGCTTATGAGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAG 2736

Search completed: March 19, 2005, 19:55:55
Job time : 1181.22 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 03:43:36 ; Search time 369.746 Seconds
(without alignments)
9983.718 Million cell updates/sec

Title: US-10-660-763-1_COPY_72_2327

Perfect score: 2256
Sequence: 1 atggggctctctctcgcgact.....gcacccgaagcgcgcgcg 2256

Scoring table:
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgnt2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgnt2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgnt2_6/prodata/1/ina/5A_COMB.seq: *
4: /cgnt2_6/prodata/1/ina/5B_COMB.seq: *
5: /cgnt2_6/prodata/1/ina/5A_COMB.seq: *
6: /cgnt2_6/prodata/1/ina/5B_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	100.0	2674	3	US-09-817-180-1
2	2256	100.0	2674	4	US-10-003-295-1
3	296.2	13.1	361	3	US-09-387-212-5
4	296.2	13.1	361	3	US-09-387-212-5
5	227.2	10.1	15297	3	US-09-948-802-5
6	227.2	10.1	15297	3	US-09-948-802-5
7	213	9.4	19152	4	US-10-003-295-3
8	213	9.4	19152	4	US-09-949-016-12110
9	198	8.8	3370	4	US-09-949-016-15795
10	185.4	8.2	3921	4	US-09-814-915A-82
11	185.4	8.2	3921	4	US-09-949-016-628
12	174.6	7.7	3386	4	US-09-949-016-2120
13	174.6	7.7	3386	4	US-09-949-016-5662
14	174.6	7.7	3416	2	US-08-357-642A-2
15	174.6	7.7	3416	2	US-08-460-626-2
16	174.6	7.7	4089	4	US-09-016-434-1483
17	170.4	7.6	1611	1	US-09-300-958A-13
18	170.4	7.6	1611	1	US-07-820-011A-3
19	170.4	7.6	1611	4	US-09-860-473-3
20	170.4	7.6	1611	4	US-09-444-711A-1
21	170.4	7.6	1611	4	US-09-444-711A-3
22	170.4	7.6	1611	5	PCT-US93-00445-3
23	169.4	7.5	2455	4	US-09-949-016-4411
24	169.4	7.5	3623	4	US-08-306-691B-35
25	169.4	7.5	5434	4	US-09-949-016-927
26	169.4	7.5	5763	4	US-09-949-016-3759
27	169	7.5	3018	4	US-09-949-016-1097
					Sequence 4749, Ap

28	169	7.5	4871	4	US-09-799-451-448	Sequence 448, Ap
29	168.2	7.5	2574	3	US-09-142-529-2	Sequence 2, Ap11
30	168.2	7.5	2574	4	US-10-045-428A-2	Sequence 2, Ap11
31	165.8	7.3	738	2	US-08-604-989A-8	Sequence 8, Ap11
32	165.8	7.3	1398	2	US-08-604-989A-9	Sequence 9, Ap11
33	165.8	7.3	1521	2	US-08-604-989A-10	Sequence 10, Ap11
34	165.8	7.3	1713	3	US-09-741-154-1	Sequence 1, Ap11
35	165.8	7.3	1942	2	US-08-604-989A-11	Sequence 11, Ap11
36	165.8	7.3	2000	3	US-08-426-509A-1	Sequence 1, Ap11
37	165.8	7.3	2000	4	US-08-232-545-1	Sequence 1, Ap11
38	163.6	7.3	2000	5	PCT-US95-05008-1	Sequence 1, Ap11
39	163.6	7.3	2449	4	US-09-949-016-1248	Sequence 1248, Ap
40	163.6	7.3	5993	3	US-09-383-630-1	Sequence 1, Ap11
41	163.6	7.3	5993	3	US-09-383-630-2	Sequence 2, Ap11
42	162.6	7.2	1467	4	US-09-579-182-2	Sequence 2, Ap11
43	162.2	7.2	2440	1	US-08-160-861-2	Sequence 2, Ap11
44	162.2	7.2	2442	1	US-08-542-363-3	Sequence 3, Ap11
45	162.2	7.2	2442	3	US-09-100-089-3	Sequence 3, Ap11

ALIGNMENTS

RESULT 1									
US-09-817-180-1									
Sequence 1, Application US/09817180									
Patent No. 6340584									
GENERAL INFORMATION:									
APPLICANT: GAN, Weinlu et al.									
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC									
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES									
FILE OF INVENTION: THREEOF									
FILE REFERENCE: CL001183									
CURRENT APPLICATION NUMBER: US/09/817,180									
CURRENT FILING DATE: 2001-03-27									
NUMBER OF SEQ ID NOS: 4									
SOFTWARE: PaateSeq for Windows Version 4.0									
SEQ ID NO 1									
LENGTH: 2674									
TYPE: DNA									
ORGANISM: Human									
US-09-817-180-1									
Query Match									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGGGCTTCCTCTGAGCGTGCAGCGCCGAGGCGGCTCTGAGCAATGACAG	60						
DB	72	ATGGGCTTCCTCTGAGCGTGCAGCGCCGAGGCGGCTCTGAGCAATGACAG	131						
QY	61	GAGGCGGCTTCCTCTGAGCGTGCAGCGCCGAGGCGGCTCTGAGCAATGACAG	120						
DB	132	GAGGCGGCTTCCTCTGAGCGTGCAGCGCCGAGGCGGCTCTGAGCAATGACAG	191						
QY	121	GACGAGGATGACGAGATGCTTCAACATGTCCTTCAAGACAGTGGGCGGACAGC	180						
DB	192	GACGAGGATGACGAGATGCTTCAACATGTCCTTCAAGACAGTGGGCGGACAGC	251						
QY	181	CGGCGCATGACGCGCTTCAAGCGCCCATGATGCTCTGGGCTGAGATCAACCAACT	240						
DB	252	CGGCGCATGACGCGCTTCAAGCGCCCATGATGCTCTGGGCTGAGATCAACCAACT	311						
QY	241	GAGGCGCTGAGCGCTTGTGCGGACGACGAGAGATCTGAATCAAGGCCCTGAGC	300						
DB	312	GAGGCGCTGAGCGCTTGTGCGGACGACGAGAGATCTGAATCAAGGCCCTGAGC	371						
QY	301	AAGCTAGGCTGCTCATCCGGAACGCGACGAGCTTCAAGATCAACGAGCGAGTGG	360						
DB	372	AAGCTAGGCTGCTCATCCGGAACGCGACGAGCTTCAAGATCAACGAGCGAGTGG	431						
QY	361	CAGCAGCTCAGAGGAGCTCAACCAAGCCAGGCGAGGACATGAGAGTGAAGAGC	420						

432 CAGCAGCTGAGAGAGCTCACCAAGACCACAGCCAGGACATGTAGAAAGCTGAAGAC 491
421 CAGTACCGAGCTTGGCAGCGGGACAGTGGCCAAAGCCAAAGCAAGTACCAAGAGGCGAC 480
492 CAGTACCGAGCTTGGCAGCGGGACAGTGGCCAAAGCCAAAGCAAGTACCAAGAGGCGAC 551
481 AAAGACAAAGACCGGTGACAAAGCCAAAGCAAGTATGTGGCAGCCTGTGGAAGCTTT 540
552 AAAGACAAAGACCGGTGACAAAGCCAAAGCAAGTATGTGGCAGCCTGTGGAAGCTTT 611
541 GCTCAGCAGACCGGTATGTGTGGCGGTGGCGGTGGCGAGCTTACCAACAGCAGCAG 600
612 GCTCAGCAGACCGGTATGTGTGGCGGTGGCGGTGGCGAGCTTACCAACAGCAGCAG 671
601 CAGCAGCTCTGTGGCGGTGGCGGTGGCGGTGGCGAGCTTACCAACAGCAGCAG 660
672 CAGCAGCTCTGTGGCGGTGGCGGTGGCGGTGGCGAGCTTACCAACAGCAGCAG 731
661 TGCATCTGAAAGAGATCTCTGAGAAATACCTGGAAGATTAGCAGCCTGTGAGATGAG 720
732 TGCATCTGAAAGAGATCTCTGAGAAATACCTGGAAGATTAGCAGCCTGTGAGATGAG 791
721 GTGTGGCCATTCACCGGAGATGAGCTGAGCTGTGGCGATCCAGCTGAGGCTGAG 780
792 GTGTGGCCATTCACCGGAGATGAGCTGAGCTGTGGCGATCCAGCTGAGGCTGAG 851
781 TACCAAGCTTCTGTGAGCAGTATGGGTCCGCACTGAGCCTCCACCTGTGTCAAGTT 840
852 TACCAAGCTTCTGTGAGCAGTATGGGTCCGCACTGAGCCTCCACCTGTGTCAAGTT 911
841 GATGAGTCACTGTGAGAGAGGTGAAACGCTGGAAGCTTGGAGAGCTTCAAGTGAAG 900
912 GATGAGTCACTGTGAGAGAGGTGAAACGCTGGAAGCTTGGAGAGCTTCAAGTGAAG 971
901 CTGACTGTGAGAGAGGTGAGCAACGCTGAGCCTTCAAGTGAAGCTTGTGGCC 960
972 CTGACTGTGAGAGAGGTGAGCAACGCTGAGCCTTCAAGTGAAGCTTGTGGCC 1031
961 ACCGAGATGTGTTTACGCGGCGAGAGATGGTTTACCAAGCTTGAACAGAGCTTCCGGAAT 1020
1032 ACCGAGATGTGTTTACGCGGCGAGAGATGGTTTACCAAGCTTGAACAGAGCTTCCGGAAT 1091
1021 GAAGAGAGAGACACCAACCCCGGAGAGGAGTCAAGCTGTGGGCAAGAGCAAGTCTG 1080
1092 GAAGAGAGAGACACCAACCCCGGAGAGGAGTCAAGCTGTGGGCAAGAGCAAGTCTG 1151
1081 CAAGAGACACTGAGAGGAGTGAAGTGAAGCTGTGAGCCAGGCAAGCTGAGGCTCAG 1140
1152 CAAGAGACACTGAGAGGAGTGAAGTGAAGCTGTGAGCCAGGCAAGCTGAGGCTCAG 1211
1141 CAGGAGTGTGTGAGACCAAGCTGTGAGCACTTGGGCGGCGAGCCCGCTGTGTG 1200
1212 CAGGAGTGTGTGAGACCAAGCTGTGAGCACTTGGGCGGCGAGCCCGCTGTGTG 1271
1201 CTCTCAGAGATGACGCGCACTCAGCTGCTCTCCGAGAGAGAGAGAGGAGGAG 1260
1272 CTCTCAGAGATGACGCGCACTCAGCTGCTCTCCGAGAGAGAGAGAGGAGGAG 1331
1261 ACAACCAACGCTGAGATCTTAAAGCCACATCTCAGAAATCTTCCGCCCAAGTTCTG 1320
1332 ACAACCAACGCTGAGATCTTAAAGCCACATCTCAGAAATCTTCCGCCCAAGTTCTG 1391
1321 AACCTGATCCGATGGAAGGAGGAGTCTTCAAGCACTTCTTGTCTCATCCAGCACTA 1380
1392 AACCTGATCCGATGGAAGGAGGAGTCTTCAAGCACTTCTTGTCTCATCCAGCACTA 1451
1381 CTGAGACCCAGAGCCCTCACCAGAAAGATGTGTCTCTGCAAGAGGCTGTGCC 1440
1452 CTGAGACCCAGAGCCCTCACCAGAAAGATGTGTGTCTCTGCAAGAGGCTGTGCC 1511
1441 AAGGACAGTGGGTGTGAACCATGAGGACCTGTGTGTGGGTGAGCAGATTTGACGAGG 1500
1512 AAGGACAGTGGGTGTGAACCATGAGGACCTGTGTGTGGGTGAGCAGATTTGACGAGG 1571

1501 AACTTTGGGAGATGTTTACGCGGAGCGCTGCGAGCCGACCAACACCCTGTGGCGTGAAG 1560
1572 AACTTTGGGAGATGTTTACGCGGAGCGCTGCGAGCCGACCAACACCCTGTGGCGTGAAG 1631
1561 TCTTGTGAGAGACGCTCCACCTTCAAGGCGCAAGTTTCTACAGGAACGAGATC 1620
1632 TCTTGTGAGAGACGCTCCACCTTCAAGGCGCAAGTTTCTACAGGAACGAGATC 1691
1621 CTGAAGCAGTACAGCCACCCCAACATCGTGGCTTCAATTGTGTGTGACCCAGAGAGC 1680
1692 CTGAAGCAGTACAGCCACCCCAACATCGTGGCTTCAATTGTGTGTGACCCAGAGAGC 1751
1681 CCCATCTACATGATGATGAGCTTGTGAGAGGAGGAGGAGCTTCTCTGACCTTCTCCGAC 1740
1752 CCCATCTACATGATGATGAGCTTGTGAGAGGAGGAGGAGCTTCTCTGACCTTCTCCGAC 1811
1741 GAGGAGGCGCGCTGCGAGTGAAGACTGTCTGAGATGATGATGATGATGATGATGATG 1800
1812 GAGGAGGCGCGCTGCGAGTGAAGACTGTCTGAGATGATGATGATGATGATGATGATG 1871
1801 ATGAGTACCTGAGAGCAAGTGTGATCCACCGGAGACTGGCTGTGGAACTTGGCTG 1860
1872 ATGAGTACCTGAGAGCAAGTGTGATCCACCGGAGACTGGCTGTGGAACTTGGCTG 1931
1861 GTGACAGAGAGATGTCTGAAAGATCAGTGAATTTGGAGTGTCCGAGAGAGAGCGGAT 1920
1932 GTGACAGAGAGATGTCTGAAAGATCAGTGAATTTGGAGTGTCCGAGAGAGAGCGGAT 1991
1921 GGGGTCTATGACGCTCAGAGGAGGCTTCAAGCAAGTCTCCGTAAGTGAACCGCAGCTGAG 1980
1992 GGGGTCTATGACGCTCAGAGGAGGCTTCAAGCAAGTCTCCGTAAGTGAACCGCAGCTGAG 2051
1981 GCCCTTAATACGCGCGCTACTCTCCGAACCGAGCTGTGAGCTTTGGCAATCTTGGCTC 2040
2052 GCCCTTAATACGCGCGCTACTCTCCGAACCGAGCTGTGAGCTTTGGCAATCTTGGCTC 2111
2041 TGGGAGACCTTCAAGCTGAGGAGGCTCCCTCTATCCCAACCTGAGCAATCAGCAGACCG 2100
2112 TGGGAGACCTTCAAGCTGAGGAGGCTCCCTCTATCCCAACCTGAGCAATCAGCAGACCG 2171
2101 GAGTTTGTGAGAGAGGAGGAGGCTGTGCGCTGCCAGAGCTGTGTCTGATGCGGTTC 2160
2172 GAGTTTGTGAGAGAGGAGGAGGCTGTGCGCTGCCAGAGCTGTGTCTGATGCGGTTC 2231
2161 AGGCTCATGAGACAGTGTGTGGCTTATGAGCTGTGGGCAAGGCGGCTTTCAGCAATC 2220
2232 AGGCTCATGAGACAGTGTGTGGCTTATGAGCTGTGGGCAAGGCGGCTTTCAGCAATC 2291
2221 TACAGAGACTGAGAGCATCCGAAAGCGGAGATCGG 2256
2292 TACAGAGACTGAGAGCATCCGAAAGCGGAGATCGG 2327

RESULT 2
US-10-003-295-1
; Sequence 1, Application US/10003295
; Patent No. 6686187
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001183DIV
; CURRENT APPLICATION NUMBER: US/10/003,295
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2674
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-1

applied

Query Match 100.0%; Score 2256; DB 4; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCTTCTCTAGCTGTGACAGCCCGGAGGCGAGGCTCTGACGAATGACG 60
DB 72 ATGGGCTTCTCTAGCTGTGACAGCCCGGAGGCGAGGCTCTGACGAATGACG 131
QY 61 GAGGCGAGCTTGTGTACTGAGGGGATGAGAAATGATGAGCCGAGGCTCAAGAT 120
DB 132 GAGGCGAGCTTGTGTACTGAGGGGATGAGAAATGATGAGCCGAGGCTCAAGAT 191
QY 121 GACAGGAGATATGACAGACTGTCTTCAACCAATGTCCCTGACAGACAGTGGGGCCAGAGC 180
DB 192 GACAGGAGATATGACAGACTGTCTTCAACCAATGTCCCTGACAGACAGTGGGGCCAGAGC 251
QY 181 GGGGCGATCAGCCCTGACAGCCCATCAGTCACTGAGGCTGAGATCAACAGCAAACT 240
DB 252 GGGGCGATCAGCCCTGACAGCCCATCAGTCACTGAGGCTGAGATCAACAGCAAACT 311
QY 241 GAGGGCTGAGCCGCTTGTCTGCGGACAGCAGAGAGATCTGAATCAAGGCCCTGAGC 300
DB 312 GAGGGCTGAGCCGCTTGTCTGCGGACAGCAGAGAGATCTGAATCAAGGCCCTGAGC 371
QY 301 AGCTGAGCTGCTCATCCGGGAAACGCGACAGACTTGGCAAGCTTACAGCGACAGTGG 360
DB 372 AGCTGAGCTGCTCATCCGGGAAACGCGACAGACTTGGCAAGCTTACAGCGACAGTGG 431
QY 361 CAGCAGCTGACAGAGAGCTCACCAAGACCACAGCAGAGCAATTGAGAACTGAAAGC 420
DB 432 CAGCAGCTGACAGAGAGCTCACCAAGACCACAGCAGAGCAATTGAGAACTGAAAGC 491
QY 421 CAGTACGAGCTGTGACACGGGACAGTGCCCAAGCCAAAGCCAAATACAGAGGCCAGC 480
DB 492 CAGTACGAGCTGTGACACGGGACAGTGCCCAAGCCAAAGCCAAATACAGAGGCCAGC 551
QY 481 AAAGACAGAGACCTGTACAAAGGCCAAGAGCAATATGTGCGACCTGTGGAAGCTTTT 540
DB 552 AAAGACAGAGACCTGTACAAAGGCCAAGAGCAATATGTGCGACCTGTGGAAGCTTTT 611
QY 541 GCTCACACAAACCGCTATGTGCTGGGGGAGTGGGGCTGCGAGCTTACCAACAGCAACAC 600
DB 612 GCTCACACAAACCGCTATGTGCTGGGGGAGTGGGGCTGCGAGCTTACCAACAGCAACAC 671
QY 601 CACCAAGCTCTGCTGCGCCGCGCTGCTGCGGCTCACTGACAGACCTGACAGAGATGCT 660
DB 672 CACCAAGCTCTGCTGCGCCGCGCTGCTGCGGCTCACTGACAGACCTGACAGAGATGCT 731
QY 661 TGCATCTTGAAAGAGATCTGCGAAGATCTGAGAAATAGCAAGCTGTGCAAGATGAG 720
DB 732 TGCATCTTGAAAGAGATCTGCGAAGATCTGAGAAATAGCAAGCTGTGCAAGATGAG 791
QY 721 TGGTGGCCCAATTCAGCGGGAGATGAGCTGACAGCTGCGCCGCACTCAAGCTTGAAGCTGAG 780
DB 792 TGGTGGCCCAATTCAGCGGGAGATGAGCTGACAGCTGCGCCGCACTCAAGCTTGAAGCTGAG 851
QY 781 TACCAAGGCTTCTGCGACAGTATGAGGTCGCACTGACGCTCCACCTGTGCACTTC 840
DB 852 TACCAAGGCTTCTGCGACAGTATGAGGTCGCACTGACGCTCCACCTGTGCACTTC 911
QY 841 GATGAGTCACTGCTTGAAGAGGGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGAAG 900
DB 912 GATGAGTCACTGCTTGAAGAGGGTGAACCGCTGAGAGCTTGGGAGCTTCAAGTGAAG 971
QY 901 CTGACTGTGAGAGCGGTGACAGCAGCTGACCTCAAGTGAAGATGAGCTGCGTGGGCC 960
DB 972 CTGACTGTGAGAGCGGTGACAGCAGCTGACCTCAAGTGAAGATGAGCTGCGTGGGCC 1031
QY 961 ACCGAGATGATGTTTACAGCGCGAGAGATGTTTACAGAGCTGCAACAGAGCTCCGGAAT 1020
DB 1032 ACCGAGATGATGTTTACAGCGCGAGAGATGTTTACAGAGCTGCAACAGAGCTCCGGAAT 1091

QY 1021 GAAAGAGAGAACACCCACCCCGGGAGCGGGTGCAGCTGTGGGCAAGAGGCAAGTGTG 1080
DB 1092 GAAAGAGAGAACACCCACCCCGGGAGCGGGTGCAGCTGTGGGCAAGAGGCAAGTGTG 1151
QY 1081 CAAGAAGCATGTGAGGGGCTGCAAGTATGAGCTGTGACAGCCAGGCAAGCTGACAGGCCAG 1140
DB 1152 CAAGAAGCATGTGAGGGGCTGCAAGTATGAGCTGTGACAGCCAGGCAAGCTGACAGGCCAG 1211
QY 1141 CAGAGTGTCTGACAGCAACAGCTGAGACCTTGGGCTCCCGGAGAGGCCCTGTGTCTG 1200
DB 1212 CAGAGTGTCTGACAGCAACAGCTGAGACCTTGGGCTCCCGGAGAGGCCCTGTGTCTG 1271
QY 1201 CTCTGCAAGATGACCGGCACTCCAGTGTCTTGGAGACAGAGGAGAGAGGGGGAAAG 1260
DB 1272 CTCTGCAAGATGACCGGCACTCCAGTGTCTTGGAGACAGAGGAGAGGGGGAAAG 1331
QY 1261 ACACCCAGCTGAGATCTTAAAGACCAATCTCAGAACTTCCGCCCAAGTCTG 1320
DB 1332 ACACCCAGCTGAGATCTTAAAGACCAATCTCAGAACTTCCGCCCAAGTCTG 1391
QY 1321 AACCTGTACCTGAGAGAGGAGGCTTCTTACGATCTTGTCTCATGACACTTA 1380
DB 1392 AACCTGTACCTGAGAGAGGAGGCTTCTTACGATCTTGTGTCTCATGACACTTA 1451
QY 1381 CTGAGACCCAGAGCCCTCACCAAGAGAGTGTGTCTCTGACAGAGGCTGTGCC 1440
DB 1452 CTGAGACCCAGAGCCCTCACCAAGAGAGTGTGTCTCTGACAGAGGCTGTGCC 1511
QY 1441 AAGGACAAATGGGTGTGAACCAATGAGACCTGTGTTGGGTGAGAGATGGAACGGGG 1500
DB 1512 AAGGACAAATGGGTGTGAACCAATGAGACCTGTGTTGGGTGAGAGATGGAACGGGG 1571
QY 1501 AACTTTGGGAAAGTGTTCAGCGGAGCGCTTGGAGCCGACCAACCTGTGTGGTGAAG 1560
DB 1572 AACTTTGGGAAAGTGTTCAGCGGAGCGCTTGGAGCCGACCAACCTGTGTGGTGAAG 1631
QY 1561 TCTTGTGAGAGAGCGTCCCACTGACCTCAAGGGCAAGTTCTACAGAAAGGAGGATC 1620
DB 1632 TCTTGTGAGAGAGCGTCCCACTGACCTCAAGGGCAAGTTCTACAGAAAGGAGGATC 1691
QY 1621 CTGAAACAGTATACAGCAACCCCAACATGCTGTGCTCATTTGATGTCTGCAACCCAGAAAGCAG 1680
DB 1692 CTGAAACAGTATACAGCAACCCCAACATGCTGTGCTCATTTGATGTCTGCAACCCAGAAAGCAG 1751
QY 1681 CCCATCTACATGCTCATGAGACTTGTGACAGGGGGGAGACTTCTGACCTTCCGACAG 1740
DB 1752 CCCATCTACATGCTCATGAGACTTGTGACAGGGGGGAGACTTCTGACCTTCCGACAG 1811
QY 1741 GAGGGGGCCCGCTGCGGGTGAAGATCTGTGCTGCANAATGTTGGGGGATGCAAGCTGTGGC 1800
DB 1812 GAGGGGGCCCGCTGCGGGTGAAGATCTGTGCTGCANAATGTTGGGGGATGCAAGCTGTGGC 1871
QY 1801 ATGAGTATCTGAGAGAGAGTGTGCTGATCAACCGGAGACTGTGCTGGAATCTGCTG 1860
DB 1872 ATGAGTATCTGAGAGAGAGTGTGCTGATCAACCGGAGACTGTGCTGGAATCTGCTG 1931
QY 1861 GTGACAGAGAGAAATGTCTGAAAGATCAGTGAATTTGGAGATGTCGAGAGAGAGCCGAT 1920
DB 1932 GTGACAGAGAGAAATGTCTGAAAGATCAGTGAATTTGGAGATGTCGAGAGAGAGCCGAT 1991
QY 1921 GGGGTCTATGCAAGCTTCAAGGGGCTTCAAGCAAGTCCCGTGAAGTGAAGCCGACCTGAG 2051
DB 1992 GGGGTCTATGCAAGCTTCAAGGGGCTTCAAGCAAGTCCCGTGAAGTGAAGCCGACCTGAG 2051
QY 1981 GCCCTTAATCTACGGCGGCTACTCTCCGAAGGAGAGTGTGAGAGCTTGTGCACTTGTCTC 2040
DB 2052 GCCCTTAATCTACGGCGGCTACTCTCCGAAGGAGAGTGTGAGAGCTTGTGCACTTGTCTC 2111
QY 2041 TGGAGAACCTTTCAGCTGAGGGGCTTCCCTTATCCCAACTCAGCAATCAGACAGACCG 2100
DB 2112 TGGAGAACCTTTCAGCTGAGGGGCTTCCCTTATCCCAACTCAGCAATCAGACAGACCG 2171
QY 2101 GATTTGTGAGAGAGGGGGCCGTCTGCTGCGCAGAGCTGTGTCTGATGCGGTTC 2160

Accession	Sequence	Position
Db	2172 GAGTTTGTGGAGAAAGGGGGGCCGTCTGCTGCTCGCCAGAGCTGTGTGCTGATGCGCGTTC	22311
Qy	2161 AGGCTCATGAGCAGCATGCTGSGCCCTATGAGCCTGSGCAGCGGCCACAGCTTGACAGCATC	22208
Db	2232 AGGCTCATGAGCAGCATGCTGSGCCCTATGAGCCTGSGCAGCGGCCACAGCTTGACAGCATC	22311
Qy	2221 TACACGAGCGTCGACAGCATCCGAAAGGGGCGATCGG	2256
Db	2292 TACACGAGCGTCGACAGCATCCGAAAGGGGCGATCGG	2327

```

RESULT 3
US-09-387-212-5
; Sequence 5, Application US/09387212A
; Patent No. 6309849
; GENERAL INFORMATION
; APPLICANT: ROBINSON, KEITH E.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
; TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
; FILE REFERENCE: MNT-090
; CURRENT APPLICATION NUMBER: US/09/387,212A
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 361
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: All occurrences of n indicate any nucleotide
US-09-387-212-5

```

Query Match	13.1%;	Score 296.2;	DB 3;	Length 361;
Best Local Similarity	97.4%;	Pred. No. 6e-56;		
Matches 333;	Conservative	0;	Mismatches 4;	Indels 5;
				Gaps 3;

QY	1613	CGAGATCTTGAAGATGACAGCCACCCCAACATGTGCGTCTCATTTGGTGTCTGCACCC	1672
Db	20	CNAGATCTTGAAGATGACAGCCACCCCAACATGTGCGTCTCATTTGGTGTCTGCACCC	79
QY	1673	AGAGCAGCCCATCATCGTCAATGAGAGCTTGTGACAGGGGGGCGACCTTCCGACCTTCC	1732
Db	80	AGAGCAGCCCATCTTACATCGTCAATGAGAGCTTGTGACAGGGGGGCGACCTTCCGACCTTCC	139
QY	1733	TCCGACGAGAGGGGGCCCGCCCTGCGGGTGAAGA CTCTGTGCAGATGTGTGGGGGATGCAG	1792
Db	140	TCCGACGAGAGGGGGCCCGCCCTGCGGGTGAAGACTCTGTGCAGATGTGTGGGGGATGCAG	199
QY	1793	CTGCTGGCATGAGATGACTTGAGAGCAAGTGTGCATCACCGGGACCTGGCTGTCTCGGA	1852
Db	200	CTGCTGGCATGAGATGACTTGAGAGCAAGTGTGCATCACCGGGACCTGGCTGTCTCGGA	259
QY	1853	ACTGCTTGTGACAGAGAAAGATGTCTTGAAAG-ATCACTGACTTT-GGGATGTCCCGAGA	1910
Db	260	ACTGCTTGTGACAGAGAAAGATGTCTTGAAAGATCACTGACTTTGGGAGATGTCCCGAGA	319
QY	1911	GGAGCC---GATGGGCTTAATGACAGCTCAGGGGGGCTTCAG	1949
Db	320	GGAGCCCGATTGGGGGTCTTAAGCAGCTCAGGGGGGCTTCAG	361

RESULT 4
 US-09-948-802-5
 : Sequence 5, Application US/09948802
 : Patent No. 6465232
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: ROBINSON, KEITH E.
 : TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE AND
 : TITLE OF INVENTION: PHOSPHATASE HOMOLOGUES AND USES THEREFOR
 : FILE REFERENCE: NMI-090
 : CURRENT APPLICATION NUMBER: US/09/948, 802
 : CURRENT FILING DATE: 2001-09-07

```

: PRIOR APPLICATION NUMBER: 09/387,212
: PRIOR FILING DATE: 1999-08-31
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 5
: LENGTH: 361
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURES:
: OTHER INFORMATION: All occurrences of n indicate any nucleotide
: US-09-948-802-5

```

Query Match	13.14%	Score 296.2	DB 3	Length 361
Best Local Similarity	97.44%	Pred. No. 6e-56		
Matches 333	Conservative 0	Mismatches 4	Indels 5	Gaps 3
QY	1613	CGAGATCCTGAAGCACTACAGCCACCCCAACATCGTCGCTTCATYGTGTGTGCACCC	1672	
Db	20	CNAGATCTGAAGAGTAACAGCCACCCCAACATCGTCGCTTCATYGTGTGTGCACCC	79	
QY	1673	AGAAAGCCCCCATCTTACATGTCATGTGAGCTTTGTGCAGGGGGGGCGACTTCTTCACTTCC	1732	
Db	80	AGAAAGCCCCCATCTTACATGTCATGTGAGCTTTGTGCAGGGGGGGCGACTTCTTCACTTCC	139	
QY	1733	TCCGCAAGAGAGGGGGCCCCGCTCGGGGTGAAGACTCGTCGACAGATGTCAGGGAGTCAG	1792	
Db	140	TCCGCAAGAGAGGGGGCCCCGCTCGGGGTGAAGACTCGTCGACAGATGTCAGGGAGTCAG	199	
QY	1793	CTGCTGGCATGAGTAACCTTGAGAGCAAGTGTCTGCATCCACCGGAGCTTGCTTCGGA	1852	
Db	200	CTGCTGGCATGAGTAACCTTGAGAGCAAGTGTCTGCATCCACCGGAGCTTGCTTCGGA	259	
QY	1853	ACTGCTGTGTGAACAGAGAAAGATGTCTTGAAG-ATCAGTGACTTT-GGATGTCCGAGGA	1910	
Db	260	ACTGCTGTGTGAACAGAGAAAGATGTCTTGAAGATCAGTACTTTGGGGATGTCCGAGGA	319	
QY	1911	GGAAAGCC---GATGGGGTCTATGAGCTCAAGGGGGGCTTGA	1949	
Db	320	GGAAAGCCGATGGGGGTCTATGAGCTCAAGGGGGGCTTGA	361	

```

RESULT 5
US-09-817-180-3
; Sequence 3, Application US/09817180
; Patent No. 6340584
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CU001183
; CURRENT APPLICATION NUMBER: US/09/817,180
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Human
US-09-817-180-3

```

Query Match	10.1%	Score 227.2;	DB 3;	Length 15297;
Best Local Similarity	71.1%	Pred. No. 3.le-40;		
Matches 394;	Conservative	0;	Mismatches 8;	Indels 152; Gaps 1;
Oy	1	ATGGGCTTCCTTGTAGCTGTGCAGCCCGCAGGCCACGSGGGCTCGACGAAATGCAG	60	
Db	2563	ATGGGCTTCCTTGTAGCTGTGCAGCCCGCAGGCCACGSGGGCTCGACGAAATGCAG	26222	
Oy	61	GAGGCCGAGGCTCGTCAACTGAGAGGGCATAGAAGTGATGAGCCCGGGTCAAGAGT	120	
Db	2623	GAGGCCGAGGCTCGTCAACTGAGAGGGCATAGAAGTGATGAGCCCGGGTCAAGAGT	26622	

Sequence 15795, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FaestSeq for Windows Version 4.0
SEQ ID NO 15795
LENGTH: 19153
TYPE: DNA
ORGANISM: Human
US-09-949-016-15795

Query Match 9.4%; Score 213; DB 4; Length 19153;
Best Local Similarity 100.0%; Pred. No. 4.4e-37;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGCTTCCTCTCTAGCTGTCAGAGCCCGGAGGCGCAAGGGGTCTCTGCAAGCAATGCG 60
DB 18741 AATGGCTTCCTCTCTAGCTGTCAGAGCCCGGAGGCGCAAGGGGTCTCTGCAAGCAATGCG 18800
QY 61 GAGGCGAGCTTCTCTAGCTGTCAGAGGCGCATGAGAAAGTGATGCGCCAGCGGTCAAGAGT 120
DB 18801 GAGGCGAGCTTCTCTAGCTGTCAGAGGCGCATGAGAAAGTGATGCGCCAGCGGTCAAGAGT 18860
QY 121 GACAGGAGTATGCAAGAGCTGCTTACCAACATGTCTCTGCAAGACAGTGGGGCCAGAGC 180
DB 18861 GACAGGAGTATGCAAGAGCTGCTTACCAACATGTCTCTGCAAGACAGTGGGGCCAGAGC 18920
QY 181 CGGGCCATCAGCCCTGACAGCCCATCAGTCAG 213
DB 18921 CGGGCCATCAGCCCTGACAGCCCATCAGTCAG 18953

RESULT 9

US-09-814-915A-82
Sequence 82, Application US/09814915A
Patent No. 6750015
GENERAL INFORMATION:
APPLICANT: Horwitz, Kathryn
APPLICANT: Richer, Jennifer
TITLE OF INVENTION: Progesterone Receptor-Regulated Gene Expression and Methods Relat
TITLE OF INVENTION: Thereeto
FILE REFERENCE: 2848-39
CURRENT APPLICATION NUMBER: US/09/814,915A
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 60/214,870
PRIOR FILING DATE: 2000-06-28
NUMBER OF SEQ ID NOS: 108
SOFTWARE: Patent version 3.1
SEQ ID NO 82
LENGTH: 3370
TYPE: DNA
ORGANISM: Homo sapiens
US-09-814-915A-82

Query Match 8.8%; Score 198; DB 4; Length 3370;
Best Local Similarity 54.7%; Pred. No. 5.2e-34;
Matches 421; Conservative 0; Mismatches 340; Indels 9; Gaps 1;

QY 1451 GGGTCTGAACCATGAGAGCACTGTGTGGGTGACAGATTGACGGGGGAACCTTTGGCG 1510
DB 1967 GGGAGCTTGATCCAGCCTGTGCTGATGTGACACATGTCAATGAGAGAGAGATTGGGG 2026

QY 1511 AAGTGTTCAGCGAGCGCTGCGAGCC-----GACAAACCCCTGTGCGGTGAAGT 1561
DB 2027 AAGTGTATGAGAGGAGACCTTCAGAGCTCCAGCAGAGACTGCAAGACTGTGGCATTAAGA 2086
QY 1562 CTGTGAGAGAGCGCTCCCACTGACCTTCAGAGCCCAAGTTTCTACAGGAAGAGAGATCC 1621
DB 2087 CCTTAAACACATCTCCCAAGGTGGCACTGTGTGAATCTTCTTCAAGAGGCAATATCA 2146
QY 1622 TGAAGCATGACAGCAACCCCAATGTCGCTCTCATTTGTTGTCACCCAGAAAGCAGC 1681
DB 2147 TGGGCGAGTTTACCCACCCGCAATTTCTCATTTGGAAGGCTTCACAAAGCAAAAGC 2206
QY 1682 CCATCTACATCTGTCATGAGAGCTTGTGACAGGGGGCGACTTCTGACCTTCTCCGACCG 1741
DB 2207 CCATCTACATCTGTCATGAGAGATTTATGGAAGAGCAGCCCTGGAATGCTTCTGAGGAGC 2266
QY 1742 AAGGGGCGCGCTGCGGGTGAAGACTCTGCTGCAATGTGTGGGATGCAAGCTGTGGCA 1801
DB 2267 GGGAGGACAGCTGTGCTCCGCGAGCTAGTGGCCATCTGCAAGGCAATAGCATTTGGGA 2326
QY 1802 TGAAGTACCTGAGAGCAAGTCTGATCCACCGGAGCTGGCTGCTGGAAGTGGCTGG 1861
DB 2327 TGAATCTCTCATGATATCAATTTATGTCCACCGGAGCTTGGCTGCAAGAAACATTTGG 2386
QY 1862 TGACAGAGAAAGATGCTCTGAGATCAGTACTTGTGGATGTCTCCGAGAGAAAGCCGATG 1921
DB 2387 TGATATCAAAACCTGTGCTGCAAGGTGTCTGACTTTGGCCCTGACTGCTCTGAGATGACT 2446
QY 1922 GGGTCTATGAGCCTCAGGGGGGCTTCAGCAAGTCCCTGTAAGTGAAGCCGACCTGAGG 1981
DB 2447 TTGATGCGCAATATGCAAAACCCAGGAGGAAAGTCTCTTCCGTTGAGACAGCCCTGAAAG 2506
QY 1982 CCTTAACTACGGCGCTACTCTCCGAAAGGAGAGTGTGAGCTTTGGCATTTGCTCT 2041
DB 2507 CATTTGCCATGATGATCTTTCACCAAGCCAGGATGTGTGAGCTTTGGGATTTGTATGT 2566
QY 2042 GGGAGACTTTCAGCCTGCGGGGCTCCCTTATCCCAACTTCAGCAATTCAGACAGACGCG 2101
DB 2567 GGGAGGTGTGAGCTTTGGGGCAAGCTTTATGGGAGATGAGCAATTCAGAGAGTTATGA 2626
QY 2102 AGTTTGTGAGAAAGGGGGGCGCTGCTCCCTGCCAAGCTGTCTGTATGCTGCTTCA 2161
DB 2627 AGAGCATTTGAGATGGGTACCGGTTGCCCTTCCTGTGACATGCGCTTGCCCTCTGTATG 2686
QY 2162 GGCTCATGAGAGAGTCTGGGCTTATGAGCTGTGGCAGCGGCCACCTTC 2211
DB 2687 AGCTCATGAAAGACTGTGGGCAATATGACCTGTGCCCGGCCACACTTC 2736

RESULT 10

US-09-949-016-628
Sequence 628, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FaestSeq for Windows Version 4.0
SEQ ID NO 628
LENGTH: 3921
TYPE: DNA
ORGANISM: Human

US-09-949-016-628

Query Match	8.2%	Score 185.4;	DB 4;	Length 3921;
Best Local Similarity	55.1%	Pred. No. 3.3e-31;		
Matches 385; Conservative	0;	Mismatches 311;	Indels 3;	Gaps 1;

Qy	1546	CTGTGTGCGGGAAGTCTTGTTCAGAGAGCGCTCCCACTGCATCTCAGAGCCAAATTTCTA	1605
Db	2037	CCGGTGGGCATTCMAGACGCTGAAGCCGGCTACACAGAAAGCAGAGTGGACTTTCCCTC	2099
Qy	1606	CAGGAAAGCAGAGATCTCTGAAGCAGTACAGCCACCCCAACATCTGTCTCATTTGGTGT	1665
Db	2097	GGCAGAGCGCGGCATTCATGGCGCAGTTGACCAACCAATCATCTCCGCTTGAAGGCGTTC	2156
Qy	1666	TGCACCCGGAAGCAGCCCATCTACATCTGATCTGAGCTTTGTGTGACAGGGAGGACATTTCCGT	1726
Db	2157	ATCTTCMAAATTCMAGCCCATGATGATCATCTAGTGTACATGAGAAATGGGGCCCTGGAC	2216
Qy	1726	ACCTTTCTCCGCACGAGAGGAGGCGCCGCTCGGGGTGAAGAATCTCTGTGCAGATGGTGGGG	1786
Db	2217	AAGTTTCTCTTGGGAGAAAGATGGCGAGTTTCAGCGTCTGCAGCTGTGTGGGCATCTCTGGG	2277
Qy	1786	GATCAGCTGTCTGGCATGAGATACCTGGAGAGCAAGTCTGCATCACCGGGACCTGGCT	1846
Db	2277	GGCATTCGAGCTTGGCATTAAGTACTTGGCCACATGAATATATGTACACCGTGAACCTGGCT	2337
Qy	1846	GCTTGGAATCTGCTGTGTGACAGAGAAAGATGTCTCTGAAGATCAGTGAATTTGGAGTGTCC	1906
Db	2337	GCCGCGAACAATCTCTGTCAACAGCAACTGTGGTCTGACAGAGGTGTGACTTTGGGCTGTCC	2399
Qy	1906	CGAAGAGAAAGCCGATGGGGTCTATGTCAAGCTT---CAGGGGGCTTACAGACAATCCCCGGT	1966
Db	2397	CGCGTGTGTGAGGAGCAGACCCGAGGCGCACTTACACACACAGTGGGGGGAAGATCCCAATC	2456
Qy	1963	AAGTGTGACCGGACCTGAGAGCCCTTAACTACAGCGCGCTACTCTCTCGAAAGCGACGTGTGG	2022
Db	2457	CGCTGTGAGCCGCGCGGAGGCCATTTCTTACCGGAAGTTCACTCTGTGCAGAGGACGTGTGG	2516
Qy	2023	AGCTTTGGCATCTTGTCTGTGGAGACCTTCAAGCTGTGGGGGCTCCCCCTATTCCAACTC	2083
Db	2517	AGCTTTGGCATATTGTCAATGTGGAGAGGTATGTACTATGTGCGAGCGGCGCTTACTGGAGATTG	2576
Qy	2083	AGCAATCAGCAGACACGGGAGTTTGTGAGAGAGGGGGGCGGTCTGCGCTGCCCAAGAGCTC	2144
Db	2577	TTCAAACACGAGGTGATGAAGCCATCAATGTGGCTTCCGGCTCCCAACCACTATGAGAC	2633
Qy	2143	TGTCTGTATGTCGCTGTTCAGGCTCAATGAGCAGATGTGTGGGCTTATGAGCTTGGAGCGG	2203
Db	2637	TGCCCCCTCCGCAATCTTACAGCTCATGATGACAGTGTGGACAGAGAGCGTGTCCGCGC	2696
Qy	2203	CCGAGCTTACACACATCTACAGGAGGTGCGAGACATC	2241
Db	2697	CCCAAGTTCGTCATCTGACATCTCTGGACAAAGCTTC	2735

RESULT 11
US-09-949-016-2120
; Sequence 2120, Application US/09949016

```
; SOFTWARE: FastSEQ for Windows Version 4.0.0
; SEQ ID NO. 3130
```

```

; SEQ ID NO 2120
; LENGTH: 3921
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2120

```

Query Match	8.2%	Score 185.4;	DB 4;	Length 3921;
Best Local Similarity	55.1%;	Pred. No. 3.3e-31;		
Matches 385;	Conservative 0;	Mismatches 311;	Indels 3;	Gaps 1;

QY 1546 CTGTGGGGGTGAAGTCTTGTGAGAGACGGCTCCACCTGACTGACCTGAAGGCTCAAGTTTCTA 1605
 Db 2037 CCGGTGGCATCAAGACCGCTGAAGCGGGCTACACAGAGAAACAGAGAGTGGACTTCCCTC 2096
 QY 1606 CAGGAAAGCAGAGATCTCTAGACGATACAGCCACCCCAATGTGTGCTCATTTGGTGTTC 1665
 Db 2097 GGCACAGCGCGGCATCATATGGCGCAGTTTCAGCCACCAACATATATCCGCTTAAGGGCGTTC 2156
 QY 1666 TGCAACCCGAAGCAGCGCCATCTCACTCGTCAATGAGCTTGTGCAAGGGGGCGCACTTCTG 1725
 Db 2157 ATCTCCAAATACAAAGCCCATATGATGATCATCACTAGTACATGGAATATGGGGCCCTGGAC 2216
 QY 1726 ACCTTCTCTCGCACGAGAGGGGGCCCGCTCGGGGTGAAGACTCTGTCTGACAGATGTGGGG 1785
 Db 2217 AAGTTCTCTTGGGAGAAAGATGCGCAGATTACAGCTGTGACAGCTGGTGGGCATGTCTGG 2276
 QY 1786 GATCAGCTGTGGCATATGAGATCACTTGGAGAGCAAGTGTGATCCACCGGGGACCTGGCT 1845
 Db 2277 GGCATTCGAGCTGGCATATGAGTACCTGGCCACATGAATATATGTGACACGTGACCTGGCT 2336
 QY 1846 GCTCGGAACCTCCCTGTGTGACAGAGAAAGATGTCTTGAGATCAGTGACTTTGGGATGTCC 1905
 Db 2337 GCCGCAACATCTCTGTCAACAGCAACCTGGTCTGCAAGGTGTGTGACTTTGGCCGTGTC 2396
 QY 1906 CGAAGGAAGCCGATGGAGTCTATGCAAGCT---CAGGGGCTCTAGACAAAGTCCCGTG 1962
 Db 2397 CGCTGTGTGAGAGACGACC CGAGGCCACTACACCAACAGTGGGGGCAAGATCCCATC 2456
 QY 1963 AAGTGAACCGACCTGAGGGCCCTTAATACAGGCGCTACTCTCCGAAAGCAAGCTGTGG 2022
 Db 2457 CGCTGACCGCCCGGAGGCCATTTCTTCCAGGAAGTTCACTCTGCCAGCAGCTGTGG 2516
 QY 2023 AGCTTTGGCATCTTGTCTGTGGAGACCTTCAAGCTGGGGGCTCCCTCCATATCCCAACTC 2082
 Db 2517 AGCTTTGGCATTTGTCATATGGGAGCGTGTATGACTTATGGGAGCGGCTTACTGGGAGTTG 2576
 QY 2083 AGCAATACGACGACACCGGAGTTTGTGGAGAAAGGGGGGCGGTGTGCCCTGCCAGAGCTG 2142
 Db 2577 TCCAACACGAGGTGATTAAGCCATCAATGATGGCTTTCGGCTCCCAACCCATGAGAC 2636
 QY 2143 TGTCTGATGCGGTGTTCAGGCTCATGAGCAGAGTGTGGGCTTATGAGCTGTGGGACGG 2202
 Db 2637 TGCCCTTCGCGCATCTACCAAGTCAATGATGCAAGTGTGTCGGCAGACGAGAGGTGCCCCCGC 2696
 QY 2203 CCCAGCTTACAGACCATATACCAAGAGCTTGACAGAGATC 2241
 Db 2697 CCCAAGTTCTGTGACATGTGAGCATCTCTGACAAGCTC 2735

/ APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: CLO011307
 / CURRENT APPLICATION NUMBER: US/09/949,016
 / CURRENT FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / PRIOR FILING DATE: 2000-09-08
 / NUMBER OF SEQ ID NOS: 207012

RESULT 12
 US-09-949-016-5662
 ; Sequence 5662, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C1001307
 ; CURRENT APPLICATION NUMBER: US/09/949, 016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241, 755
 ; PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5662
LENGTH: 3386
TYPE: DNA
ORGANISM: Human
US-09-949-016-5662

Query Match 7.7%; Score 174.6; DB 4; Length 3386;
Best Local Similarity 54.7%; Pred. No. 7.5e-29;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

QY 1498 GGGAACTTTGGCGAAGTTCAGCGAGCGCTGCGAGCCGACAAACCTCTGGGGGTG 1557
DB 1386 GGGGAGGTCTATAGAGGTGTCTACAAATCAGAAAGGGAGAAATCAATGTAGCTGTC 1445
QY 1558 AAGCTTGTGAGAGAGCGTCCACCTGACCTCAAGGCCAAAGTTTCTACAGAAAGGAG 1617
DB 1446 AAGACTCTGAGAAAGAGCTGCACTTGGACACAGAGAAAGTTTCAAGCGAGGAGTG 1505
QY 1618 ATCTGAGAGAGTACAGCCCAACCATGCTGCTCATTTGGTGTCTGACCCAGAG 1677
DB 1506 ATCATAGAAAGACTCGACCAACCGCATGTGTGAAGCTGATCGGCAT---CATTTAAGAG 1562
QY 1678 CAGCCCATCTACATCTGATGAGACTTGTGAGGGGGGCACTTCTGACCTTCTCCG 1737
DB 1563 GAGCCCACTGAGATCATGATGAAATGTATCCCTATGGGAGCTGGGCCCATCTACCTGAG 1622
QY 1738 ACGGAGGGGGCGCGCTGGGGGAGAACTCTGCTGAGATGGTGGGGAGTGAAGCTGCT 1797
DB 1623 CGGAACAGAACTCCCTGAAGGTGCTCACCTGTGTGATCTGCTGCAATATGCAAA 1682
QY 1798 GGCATGAGTACCTGAGAGCAAGTGTGATCCAGCGGAGCTGTGCTCGAACTGC 1857
DB 1683 GCCATGGCTTACCTGAGAGCATCACTGGGTGACAGGGAACATGTGTCGCCGAACATC 1742
QY 1858 CTGGTACAGAGAAATGTCCTGAGAGATGATGATCTTTGGATGTCCGAGAGAAAGCC 1917
DB 1743 CTGTGCTCCCTCCCTGAGTGTGTAAGCTGGGGGACTTTGTCTTCCCGTACATTGAG 1802
QY 1918 GATGGGCTATGAGGCTGAGGGGGGCTCAGCAAGTCCCGGAGATGAGACGACCT 1977
DB 1803 GAGGAGGACTAT--TACAAAGCTCTGTGATCTGTCTCCCATCAATGATGTCCCA 1859
QY 1978 GAGGCTTACTACGCGCGCTACTCTCCGAAAGCAGGTGAGACTTTGGCATCTTG 2037
DB 1860 GAGTCCATTAACCTTCGACGCTTACAGACAGCCAGTGAACGTGTGATGTTCCGCGTGTG 1919
QY 2038 CTCTGGGAGACTTACGCTGGGGGCTCCCTCATTCCTCAACTCAGCAATCAGAGACA 2097
DB 1920 ATGTGGGAGATCTGAGCTTTGGGAGACAGCCCTTCTTCTGCTGAGAAACAGAGATGC 1979
QY 2098 CGGAGATTGAGAGAGGGGGGCGTCTGCGCTGCCAGAGCTGTCTGATAGCCGTG 2157
DB 1980 ATCGGGTGTCTGAGAGAGAGAGAGCGGCTGCCAAGCTGATCTCTGTCAACGCGTCTT 2039
QY 2158 TTACAGCTCATGAGAGAGTCTGGGCTTATGAGCTTGGGAGCGGCCAGCTTCAGC 2214
DB 2040 TATACCTCATGAGAGCGCTGTGGGACTAGAACCCAGTGAACGGGCCCGCTTCACC 2096

RESULT 13
US-08-357-642A-2
Sequence 2, Application US/08357642A
Patent No. 5837524
GENERAL INFORMATION:
APPLICANT: Sima Lev
APPLICANT: Joseph Schlessinger
TITLE OF INVENTION: PYK2 RELATED PRODUCTS

TITLE OF INVENTION: AND METHODS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,642A
FILING DATE: December 15, 1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/070
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic

Query Match 7.7%; Score 174.6; DB 2; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.5e-29;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

QY 1498 GGGAACTTTGGCGAAGTTCAGCGAGCGCTGCGAGCCGACAAACCTCTGGGGGTG 1557
DB 1416 GGGGAGGTCTATGAGAGGTGTCTACAAATCAGAAAGGGAGAAATCAATGTAGCTGTC 1475
QY 1558 AAGCTTGTGAGAGAGCGTCCCACTGACCTCAAGGCCAAAGTTTCTACAGAAAGGAG 1617
DB 1476 AAGACTTGAAGAAAGAGCTGCACTTGACAAACAGAGAAAGTTTCAATGAGCGAGAGTG 1535
QY 1618 ATCTGAGAGAGTACAGCCCAACCATGCTGCTCATTTGGTGTCTGACCCAGAAAG 1677
DB 1536 ATCATAGAAACCTTCACCAACCGCATGTGTAAGCTGATGCGCAT---CATTTAAGAG 1592
QY 1678 CAGCCCATCTACATCTGATGAGACTTGTGAGGGGGGCGACTTCTGACCTTCTCCG 1737
DB 1593 GAGCCCACTGATCATCTGAAATTTGTATCCCTATAGGGAGAGTGGGCCACTGAGAG 1652
QY 1738 ACGGAGGGGGCGCGCTGGGGTGAAGACTGTGCTGCAAGTGTGGGGAGTGAAGCTGCT 1797
DB 1653 CGGAACAGAAAGTCCCTGAAGGTGCTCACCTGTGTGATCTGATGAGATATCAAA 1712
QY 1798 GGCATGAGTACTTGAAGAGCAAGTGTGATCTACCCGGAAGCTGTGCTCGGAATGC 1857
DB 1713 GCCATGGCTTACTGAGAGAGATCAACATGCTGTGACAGAGGAACTTCTGTCCGAACTC 1772
QY 1858 CTGGTACAGAGAAAGATGCTGGAAGATCAGTGAATTTGGAGTGTCCGAGAGAAAGCC 1917
DB 1773 CTGTGCTCCCTCCCTGAGTGTGTAAGCTGGGGGACTTTGGTCTTCCCGGTACATTGAG 1832
QY 1918 GATGGGCTATGAGAGCTTCAAGGGGGCTTCAGAAAGTCCCGTGAAGTGAACCGACCT 1977

Db 1833 GAGGAGACTAT--TACAAAGCCTCTGTGATCTGCTCTCCCATCAATGAGATGCCCCA 1889
Qy 1978 GAGGCTTAACTACGAGCGCTCTCTCCGAAAGGAGAGCTGTGAGCTTTGGCATCTTG 2037
Db 1890 GAGTCCATTAACTCCGAGCGCTTACAGAGAGCGAGTGAAGTGTTCGCGGTG 1949
Qy 2038 CTCTGGAGACCTTCAGCTGTGGGGGCTCCCTTATCCCACTCAGCAATCAGCAGACA 2097
Db 1950 ATGTGGAGATCTCTGAGCTTTGGGAGCAGCCCTTCTTGTGCTGAGAACAGAGATGTC 2009
Qy 2098 CCGGAGTTTGTGAGAGGGGGCGCTGTGCGCCGAGAGCTGTCTGATGCGGTG 2157
Db 2010 ATCGGGGTCTGTGAGAAAGAGACCGGCTGCCAAGCTGATCTGTCCACCGGTCTT 2069
Qy 2158 TTCAGGCTCATGAGAGAGTGTGGGCTTATGAGCTTGGGAGCGGCGGCACTTCAAC 2214
Db 2070 TATACCTCATGACCGGCTGTGGACTAGACCCCAAGTACCGGCGGCTTCAAC 2126

RESULT 14
US-08-460-626-2

/ Sequence 2, Application US/08460626
/ Patent No. 5837815
/ GENERAL INFORMATION:
/ APPLICANT: SIMA LEV
/ APPLICANT: JOSEPH SCHLESSINGER
/ TITLE OF INVENTION: PYK-2 RELATED PRODUCTS AND
/ TITLE OF INVENTION: METHODS
/ NUMBER OF SEQUENCES: 25
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Lyon & Lyon
/ STREET: 633 West Fifth Street
/ CITY: Los Angeles
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 90071-2066
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ MEDIUM TYPE: Storage
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: IBM P.C. DOS 5.0
/ SOFTWARE: Word Perfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/460,626
/ FILING DATE: June 2, 1995
/ CLASSIFICATION: 435
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER: 08/357,642
/ FILING DATE: December 15, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Marburg, Richard J.
/ REGISTRATION NUMBER: 32,327
/ TELEPHONE: (213) 489-1600
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (213) 955-0440
/ TELEFAX: (213) 955-0440
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3416
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-460-626-2

Query Match 7.7%; Score 174.6; DB 2; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.5e-29;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;
Qy 1498 GGGAGCTTGGCGAAGTGTTCAGCGGCGCTGCGAGCGAGCAACCCCTGTGGGGGTG 1557

Db 1416 GGGAGGTCTATGAGGTGTCTACACAATTCACAAAGGGAGAAATCATGTAGCTGTC 1475
Qy 1558 AAGTCTTGTGAGAGAGCGTCCCACTGACTCTCAAGGCCAAGTTTCTACAGAGAGGAG 1617
Db 1476 AAGACCTGCAGAAAGACTGACTCTGAGCAACCAAGAGAAAGTTCATGAGCGAGGAGTG 1535
Qy 1618 ATCTGGAACAGATACAGCCACCACCAATGATGCGCTCATTTGTGTCTGCACCCAGAA 1677
Db 1536 ATCATGAAAGACCTGACCAACCGGACATGTGAGAGTATGCGCAT--CATTGAAG 1592
Qy 1678 CAGCCCATCTACATCTCATGAGAGCTTGTGACAGGGGGGAGACTTCTGACTCTCCG 1737
Db 1593 GAGCCCACTGATATCATGAAATTTATCCCTATAGGGAGAGCTGGCCACTCTGGAG 1652
Qy 1738 ACGAGGGGGCGGCTGCGGGTGAAGACTGTGCTGACAGATGATGGGGAGATGACGTCT 1797
Db 1653 CCGAACAGAACTCCCTGAAGGTGCTCACCTCTGTGCTACTCATGACGATATCAAA 1712
Qy 1798 GGCATGAGATACCTGAGAGAGAGTGTGCTGATCCACCGGGACCTGCTGCGGACCTG 1857
Db 1713 GGCATGCTTACCTGAGAGAGATCAACATGCTGACAGAGACATTTCTCGGAACATC 1772
Qy 1858 CTGTGACAGAGAGAGATGCTCTGAAGATCAGTGAATTTGGAGATGCTCCGAGAGAGCC 1917
Db 1773 CTGTGCTCTCCCTGAGTGTGTGAAGCTGGGGGACTTGTGCTTCCGGTACATTTAG 1832
Qy 1918 GATGGGTCTATGACAGCTTCAGGGGGCTTCAGACAAATGCTCCCTGAGTGAACCGACCT 1977
Db 1833 GAGGAGAGACTAT--TACAAAGCTGTGTGACTGCTGCCATCAATGAGATGTCCCA 1889
Qy 1978 GAGGCTTAACTACGCGCGCTATCTCTCCGAAAGGAGAGTGTGAGCTTTGSCATCTT 2037
Db 1890 GAGTCCATTAACTCCGAGCTTCACAGACAGCAGTGAAGTGTGCGGTGTG 1949
Qy 2038 CTCTGGAGACCTTCAGCTGTGGGGGCTCCCTTATCCCACTCAGCATCAGCAGACA 2097
Db 1950 ATGTGGAGATCTGAGCTTTGGGAAGCAGCCTTCTTGTGCTGAGAAAGATGTC 2009
Qy 2098 CCGGAGTTTGTGAGAGGGGGCGCTGTGCGCCGAGAGCTGTGCTGATGCGGTG 2157
Db 2010 ATCGGGGTCTGAGAAAGAGACCGGCTGCCAAGCTGATCTGTCCACCGGTCTT 2069
Qy 2158 TTCAGGCTCATGAGAGAGTGTGGGCTTATGAGCTTGGGAGCGGCGGCACTTCAAC 2214
Db 2070 TATACCTCATGACCGGCTGTGGACTAGACCCCAAGTACCGGCGGCTTCAAC 2126

RESULT 15
US-09-016-434-1483

/ Sequence 1483, Application US/09016434
/ Patent No. 6500938
/ GENERAL INFORMATION:
/ APPLICANT: Janice Au-Young
/ APPLICANT: Jeffrey J. Seilhamer
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
/ TITLE OF INVENTION: PATHWAY GENE EXPRESSION
/ NUMBER OF SEQUENCES: 1490
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/016,434
/ FILING DATE: HEREWITH
/ CLASSIFICATION:

Search completed: March 19, 2005, 14:24:17
Job time : 373.746 secs

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1483:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 9988304
US-09-016-434-1483

Query Match 7.7%; Score 174.6; DB 4; Length 3416;
Best Local Similarity 54.7%; Pred. No. 7.5e-29;
Matches 392; Conservative 0; Mismatches 319; Indels 6; Gaps 2;

QY 1498 GGGAACTTGGCGAAGTGTTCAGCGGACGCGTGGACCGACAAACCTGTGTGGCGGTG 1557
DB 1416 GGGAGAGTCTATGAGGTGTCTACAAATACAAAGGAGAAATCATATGACTGTC 1475
QY 1558 AAGTCTTGTGAGAGAGCTGCCAAGCTGACCTCAAGGCCAAAGTTCTACAGAAAGCGAGG 1617
DB 1476 AAGACCTGCAAGAAAGACTGCACTGTGACACAAAGGAAAGTTTCATGACGAGGCGAGTG 1535
QY 1618 ATCTGAAGAGATACAGACCAACCCCAATGTGTGCTCATTTGTGTCTGACACCAAG 1677
DB 1536 ATCATGAAGAACTCGACCAACCGCATGTGAAAGCTGATCGGCAT--CATTTGAAGAG 1592
QY 1678 CAGGCCATCTACATCGTCATGAGCTGTGACAGGGGGGCGACTTCTGACCTTCTCCGCG 1737
DB 1593 GAGGCCACTGAGATCATGAGATTGTATCCATGAGGAGCTGGGCCACTACCTGGAG 1652
QY 1738 ACGGAGGGGGCCCGCTGCGGGTGAAAGACTGTGCTGCAGATGTGGGGGATGCAAGCTGCT 1797
DB 1653 CGGAACAAAGAACTCCCTGAAGGTGCTCACCCCTGCTGTACTGACATGACAGATATGCAAA 1712
QY 1798 GGCATGAGATACCTGAGAGAGCAAGTGTGATCCACCGGAGCTGGCTCGGAATGTC 1857
DB 1713 GCCATGGCTTACCTGAGAGACATCACTGCGTGCACAGGAGCAATGCTGTCCGGAACATC 1772
QY 1858 CTGGTGACAGAGAAATGTCTCTGAAGATCAGTCACTTTGGGATGTCCCGAGAGAGAGCC 1917
DB 1773 CTGGTGCCCTCCCTGAGTGTGTAAGCTGGGGGACTTTGGTCTTTCCCGTACATTGAG 1832
QY 1918 GATGGGCTTATGCAAGCTCAAGGGGCGCTCAACAACTCCCGTGAAGTGACCGCACT 1977
DB 1833 GACGAGAGACTAT--TACAAAGCTCTGTGACTGTCTCTCCCAATCAATGATGTCCCA 1889
QY 1978 GAGGCCCTTAACTACGCGCCGCTACTCTCCGAAAGCAAGTGTGAGACTTTGGCATTTG 2037
DB 1890 GAGTCCATTAACTTCCAGCGCTTCCAGACAGCCAGTACGCTGAGATGTTCCGCTGTGC 1949
QY 2038 CTCTGGGAGACCTTCAGCTGAGGGGCGCTCCCGCTATCCCAACTCAGCAATCAGACAGA 2097
DB 1950 ATGTGGAGATCTGAGCTTTGGGAAAGACCTTTCTTGTGGCTGGAGAAACAGAGATGTC 2009
QY 2098 CGGAGATTGTGAGAAAGGGGCGCGCTGTGCTGCCAGAGCTGTGCTGTATGCCGTG 2157
DB 2010 ATCGGGGTGCTGAGAGAGAGACCGGCTGCCAAGCTGATCTGTCCACCGGTCTT 2069
QY 2158 TTCAGGCTCATGAGAGCTGTGGGCTTATGAGCTGGGCAAGCGGCCCAAGCTTCAGC 2214
DB 2070 TATACCTCATGACCCGCTGCTGGGACTACGACCCCAAGTACCGGCCCCGCTTCACC 2126

61 GAGGCCGAGCTTGTCTAAGGAGGAGATGAGAAAGTGAAGTGGCCGAGGGGTCAAGAGT 120
DB GAGGCCGAGCTTGTCTAAGGAGGAGATGAGAAAGTGAAGTGGCCGAGGGGTCAAGAGT 151
QY 121 GACAGGAGATGATGAGAGATCTGCTTCAACAATGTCTCTGAGAGACAGTGGGGGCGAGAC 180
DB 152 GACAGGAGATGATGAGAGATCTGCTTCAACAATGTCTCTGAGAGACAGTGGGGGCGAGAC 211
QY 181 GGGGCGATCAAGCCCTGAGAGCCCATCAATCAATCTCTGGGGTGAATCAACAGCCAACT 240
DB 212 GGGGCGATCAAGCCCTGAGAGCCCATCAATCAATCTCTGGGGTGAATCAACAGCCAACT 271
QY 241 GAGGGCTGAGCCGCTTGTCTGAGGAGAGAGAGAGATCTGAATCTGAAGGAGCCCTGAGC 300
DB 272 GAGGGCTGAGCCGCTTGTCTGAGGAGAGAGAGAGATCTGAATCTGAAGGAGCCCTGAGC 331
QY 301 AAGCTGAGCCCTGCTATCCGGGAAAGGAGAGAGATCTGCAAGATCTGAGAGAGAGTGG 360
DB 332 AAGCTGAGCCCTGCTATCCGGGAAAGGAGAGAGATCTGCAAGATCTGAGAGAGAGTGG 391
QY 361 CAGCACTGAGAGAGAGAGATCTCAACAGAGCCCAAGAGAGAGATCTGAAGAGAGAGAGC 420
DB 392 CAGCACTGAGAGAGAGAGATCTCAACAGAGCCCAAGAGAGAGATCTGAAGAGAGAGAGC 451
QY 421 CAGTACCGAGCTCTGAGAGAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGC 480
DB 452 CAGTACCGAGCTCTGAGAGAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGC 511
QY 481 AAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGATCTGAGAGAGAGAGAGAGAGAGC 540
DB 512 AAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGATCTGAGAGAGAGAGAGAGAGAGC 571
QY 541 GCTCAACAGC 600
DB 572 GCTCAACAGC 631
QY 601 CACCACTCTGAGC 660
DB 632 CACCACTCTGAGC 691
QY 661 TGCATCTGAAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 720
DB 692 TGCATCTGAAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 751
QY 721 TGCATCTGAAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 780
DB 752 TGCATCTGAAGAGAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 811
QY 781 TACCAAGC 840
DB 812 TACCAAGC 871
QY 841 GATGAGCTGAGC 900
DB 872 GATGAGCTGAGC 931
QY 901 CTAATCTGAGC 960
DB 932 CTAATCTGAGC 991
QY 961 ACCGAGC 1020
DB 992 ACCGAGC 1051
QY 1021 GAGC 1080
DB 1052 GAGC 1111
QY 1081 GAGC 1140
DB 1112 GAGC 1171

QY 1141 CAGAGTGTCTGAGC 1200
DB 1172 CAGAGTGTCTGAGC 1231
QY 1201 CTCTCTGAGC 1260
DB 1232 CTCTCTGAGC 1291
QY 1261 ACACCGAGC 1319
DB 1292 ACACCGAGC 1351
QY 1320 ----- 1319
DB 1352 CTCTCTGAGC 1411
QY 1320 ----- 1319
DB 1412 TACCAAGC 1471
QY 1320 ----- 1319
DB 1472 CTCTCTGAGC 1531
QY 1320 ----- 1319
DB 1532 CCGCGAGC 1591
QY 1320 -----GAGC 1366
DB 1592 CCGCGAGC 1651
QY 1367 TCATGAGC 1426
DB 1652 TCATGAGC 1711
QY 1427 ACAGC 1486
DB 1712 ACAGC 1771
QY 1487 AGATGAGC 1546
DB 1772 AGATGAGC 1831
QY 1547 TGGTGGAGC 1606
DB 1832 TGGTGGAGC 1891
QY 1607 AGC 1666
DB 1892 AGC 1951
QY 1667 GAGC 1726
DB 1952 GAGC 2011
QY 1727 CTTCTCTGAGC 1786
DB 2012 CTTCTCTGAGC 2071
QY 1787 ATGAGC 1846
DB 2072 ATGAGC 2131
QY 1847 CTCTGAGC 1906
DB 2132 CTCTGAGC 2191
QY 1907 GAGC 1966
DB 2192 GAGC 2251
QY 1967 GAGC 2026

Db		2252	GGAACCGACCTGAGGCCCTTAAC	-----	2277
Oy		2027	TTGGCATTCCTCTGGAGACTTTCAGCCTGGGGGCTCCCTCAATCCAACTCAGCA	-----	2086
Db		2278	-----	-----	2277
Oy		2087	ATCAGCAGACAGGGAGTTTGAGAGAGGGGGCCGTCTGCCTCCAGCAGAGCTGTGC	-----	2146
Db		2278	-----	GGGGGCGCTCTGCTCCGACCAGAGCTGTGTC	2308
Oy		2147	CTGATGCCGNTTCAAGGCTCATGAGCAGTCTGggcccttatgagccttgaggacgcgccca	-----	2206
Db		2309	CTGATGCCGNTTCAAGGCTCATGAGCAGTCTGggcccttatgagccttgaggacgcgccca	-----	2368
Oy		2207	GCTTCAGACCATCTTACAGAGAGCTGCGAGAGATCCGAAAGCGGATGCG	-----	2256
Db		2369	GCTTCAGACCATCTTACAGAGAGCTGCGAGAGATCCGAAAGCGGATGCG	-----	2418
RESULT 2					
CRS93957					
LOCUS					
DEFINITION			full-length cDNA clone CSODI008Yg07 of Placenta Cot 25-normalized		
ACCESSION			CR593957		
VERSION			CR593957.1		
KEYWORDS			HTC; CNSLT_cDNA.		
SOURCE			Homo sapiens		
ORGANISM			Homo sapiens		
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE			L1.W.B., Gruber,C., Jeesee,J. and Polayes,D.		
JOURNAL			Full-length cDNA libraries and normalization		
REMARK			Unpublished		
			Contact : Feng liang Email : fliang@lifetech.com URL :		
			http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		
			Faraday Avenue		
			2 (bases 1 to 1189)		
REFERENCE					
AUTHORS			Genoscope.		
TITLE			Direct Submission		
JOURNAL			Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :		
			BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr		
			- Web : www.genoscope.cns.fr)		
COMMENT			1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime		
			end enriched, double-strand cDNA was digested with Not I and cloned		
			into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library		
			was normalized. Library was constructed by Life Technologies, a		
			division of invitrogen.		
FEATURES			Location/Qualifiers		
source			1..1189		
			/organism="Homo sapiens"		
			/mol_type="mRNA"		
			/db_xref="taxon:9606"		
			/clone="CSODI008Yg07"		
			/tissue_type="Placenta		
			/plasmid="pCMVSPORT_6"		
ORIGIN					
Query Match			44.4% Score 1002; DB 3; Length 1189;		
Best Local Similarity			100.0%; Pred. No. 1.2e-200;		
Matches 1002; Conservative			0; Mismatches 0; Indels 0; Gaps 0;		
Oy		1255	GGAAGAACAACCAACCTGTGAATCTTTAAGACCAATCTCAGAAATTTCCGCCCAAG	-----	1314
Db		1	GGAAGAACAACCAACCTGTGAATCTTTAAGACCAATCTCAGAAATTTCCGCCCAAG	-----	60
Oy		1315	TTCTGCACTGTACCGACTGGAAAGGAGCTTTCTAGCATCTTTGGCTCATGCAC	-----	1374
Db		61	TTCTGCACTGTACCGACTGGAAAGGAGCTTTCTAGCATCTTTGGCTCATGCAC	-----	120
Oy		1375	CACCTACTGAGCACCCAGAGCCCCCTACACAAAGAAGTGTGTTGCTCTGCAGAGCT	-----	1433

Db	121	CACCTTCTGAGCAACCCAGACGCCCTTCAACAAAGAAAGTGGTGTGTTCTGTCGACAGGGCT	180
Qy	1435	GTGCCCCAAGGACAAAGTGGGTCTGAACCATGAGAGACCTTGGTGGGTGAGCAGATTGGA	1494
Db	181	GTGCCCAAGGACAAAGTGGGTCTGAACCATGAGAGACCTGGTGTTGGGTGAGCAATTGGA	240
Qy	1495	CGGGGAAACTTTGGCCGAAGTCTTACGGGACGCTTGGAGACCCGCAACACCCCTGGTGGC	1554
Db	241	CGGGGAAACTTTGGCCGAAGTCTTACGGGACGCTTGGAGACCCGCAACACCCCTGGTGGC	300
Qy	1555	GTGAAGTCTTGTGAGAGAGCGCTCCCACTGACCTCAAGGCAAGTTTCTACAGAAAGC	1614
Db	301	GTGAAGTCTTGTGAGAGAGCGCTCCCACTGACCTCAAGGCAAGTTTCTACAGAAAGC	360
Qy	1615	AGGATCTTGAAGCAGTACAGCCACCCCAACATCGTGGCTCATTTGGTGTGACCCAG	1674
Db	361	AGGATCTTGAAGCAGTACAGCCACCCCAACATCGTGGCTCATTTGGTGTGACCCAG	420
Qy	1675	AAGCAGCCCATCTACATCGTGCATGAGCTTGTGCAAGGGGGGCACTTCTGACCTTCTC	1734
Db	421	AAGCAGCCCATCTACATCGTGCATGAGCTTGTGCAAGGGGGGCACTTCTGACCTTCTC	480
Qy	1735	CGCAGGAGGGGGGGCGCGCTGGCGGGTGAAGA CTGCTGTCGAGATGATGGGGATGACCT	1794
Db	481	CGCAGGAGGGGGGGCGCGCTGGCGGGTGAAGA CTGCTGTCGAGATGATGGGGATGACCT	540
Qy	1795	GCTGGCATTGAGTACTCTGAGAGAGCAGTGTCTGATCCACCGGGA CTTGGCTCGAAC	1854
Db	541	GCTGGCATTGAGTACTCTGAGAGAGCAGTGTCTGATCCACCGGGA CTTGGCTCGAAC	600
Qy	1855	TGCTCTGTGACAGAGAAAGATGTCTTGAAGTCACTGACTTTGGAGTGTCCGAGAGAA	1914
Db	601	TGCTCTGTGACAGAGAAAGATGTCTTGAAGTCACTGACTTTGGAGTGTCCGAGAGAA	660
Qy	1915	GCGAGTGGGGGTCTATGACAGCTCAGGGGGGCTCAGACAAAGTCCCGGTGAATGAGCCGCA	1974
Db	661	GCGAGTGGGGGTCTATGACAGCTCAGGGGGGCTCAGACAAAGTCCCGGTGAATGAGCCGCA	720
Qy	1975	CCTGAGGCGCTTAACTACAGCGCGCTACTCTCCGAAAGCGAGCTGTGGAGCTTTGGCATC	2034
Db	721	CCTGAGGCGCTTAACTACAGCGCGCTACTCTCCGAAAGCGAGCTGTGGAGCTTTGGCATC	780
Qy	2035	TTGCTCTGAGAGACCTTCAAGCTTGGGGGCTTCCCTATCTCCAACTCAGCATCAGAG	2094
Db	781	TTGCTCTGAGAGACCTTCAAGCTTGGGGGCTTCCCTATCTCCAACTCAGCATCAGAG	840
Qy	2095	ACAGCGGAGTTTGTGAGAAAGGGGGGCGCTTGGCCCTGAGAGCTGTGCTCTGATGCC	2154
Db	841	ACAGCGGAGTTTGTGAGAAAGGGGGGCGCTTGGCCCTGAGAGCTGTGCTCTGATGCC	900
Qy	2155	GTGTTTCAAGGCTCATGAGAGCAGTGTCTGGGCTCATAGAGCTGGGCGAGCGGCCAGCTTACG	2214
Db	901	GTGTTTCAAGGCTCATGAGAGCAGTGTCTGGGCTCATAGAGCTGGGCGAGCGGCCAGCTTACG	960
Qy	2215	ACCATCTTACAGAGAGCTGACAGAGCATCCGAAAGCGGAGTCG	2256
Db	961	ACCATCTTACAGAGAGCTGACAGAGCATCCGAAAGCGGAGTCG	1002
RESULT 3			
LOCUS	BX356089	936 bp	mRNA
DEFINITION	BX356089 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA		linear
ACCESSION	BX356089		
VERSION	BX356089.2		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	1 (bases 1 to 936)		Mammalia; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Li, W.-B., Gruber, C., Jeesee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On May 5, 2003 this sequence version replaced g1:30372027.

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7663.f

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CSODI008AD04P1ac=7663.f.

FEATURES

source

Location/Qualifiers

1..936
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI008Y07"
 /clone_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 40.9%; Score 923.6; DB 5; Length 936;
 Best Local Similarity 99.8%; Pred. No. 4.1e-184;
 Matches 934; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1255 GGAGAGACACCCGCTGAGATCTTAAAGCCACATCTCAGGAATCTCCGCCCAAG 1314
 DB 1 GGAGAGACACCCGCTGAGATCTTAAAGCCACATCTCAGGAATCTCCGCCCAAG 60
 QY 1315 TTCTCGAACCTGACGAGTGAAGAGGGAAGGCTTCTGACATCTTCTGATCGAC 1374
 DB 61 TTCTCGAACCTGACGAGTGAAGAGGGAAGGCTTCTGACATCTTCTGATCGAC 120
 QY 1375 CACCTAAGACACCCAGACGCTTACCAAGAGTGTGTCTGACAGAGGCT 1434
 DB 121 CACCTAAGACACCCAGACGCTTACCAAGAGTGTGTCTGACAGAGGCT 180
 QY 1435 GTGCCCAAGACAGTGTGTCTGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1494
 DB 181 GTGCCCAAGACAGTGTGTCTGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
 QY 1495 CGGGGGAACCTTGGCGAAGTTCAGCGGAGCGCTGCGAGCCGACCAACCTCTGTGGG 1554
 DB 241 CGGGGGAACCTTGGCGAAGTTCAGCGGAGCGCTGCGAGCCGACCAACCTCTGTGGG 300
 QY 1555 GTGAAGCTTGTGAGAGAGCTCCACCTGACCTCAAGGCCAAGTTTCTACAGAGAG 1614
 DB 301 GTGAAGCTTGTGAGAGAGCTCCACCTGACCTCAAGGCCAAGTTTCTACAGAGAG 360
 QY 1615 AGGATCTCTGAAGACGACCCACCCCAATGTGTGTGTGTGTGTGTGTGTGTGTGT 1674
 DB 361 AGGATCTCTGAAGACGACCCACCCCAATGTGTGTGTGTGTGTGTGTGTGTGTGT 420
 QY 1675 AAGCAGCCATCTACATGTCATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1734
 DB 421 AAGCAGCCATCTACATGTCATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
 QY 1735 CGCAGCGAGGGGCGCGCTGCGGGTGAAGACTCTGCTGAGATGTGTGGGAGTGA 1794
 DB 481 CGCAGCGAGGGGCGCGCTGCGGGTGAAGACTCTGCTGAGATGTGTGGGAGTGA 540
 QY 1795 GCTGGCATGAGTACTGTGAGAGCAAGTGTGATCCACGGGACCTGTGCTCGGAG 1854
 DB 541 GCTGGCATGAGTACTGTGAGAGCAAGTGTGATCCACGGGACCTGTGCTCGGAG 600

QY 1855 TGCCTGTGACAGAGAGAAATGCTCCTGAAGATCAGTACTTTGGATGTCCGAGAGAA 1914
 DB 601 TGCCTGTGACAGAGAGAAATGCTCCTGAAGATCAGTACTTTGGATGTCCGAGAGAA 660
 QY 1915 GCCGATGGGGTTTATGACCTTCAGGGGGCTTCAGACAGTCCCTGAAGTGAACCGA 1974
 DB 661 GCCGATGGGGTTTATGACCTTCAGGGGGCTTCAGACAGTCCCTGAAGTGAACCGA 720
 QY 1975 CCTGAGGGCCCTTACATGACGGCGCTACTCTCCGAAGGAGGTGTGAGCTTTGGCATC 2034
 DB 721 CCTGAGGGCCCTTACATGACGGCGCTACTCTCCGAAGGAGGTGTGAGCTTTGGCATC 780
 QY 2035 TTGCTCTGGAGACCTTCAGCCT-GGGGGCTTCCCTTATCCCACTCAGCAATCAGCA 2093
 DB 781 TTGCTCTGGAGACCTTCAGCCTGCGGGGGCTTCCCTTATCCCACTCAGCAATCAGCA 840
 QY 2094 GACACGGAGATTGTGAGAGAGGGGGCGCTGTGCTGCCAGAGCTGTGCTGATGC 2153
 DB 841 GACACGGAGATTGTGAGAGAGGGGGCGCTGTGCTGCCAGAGCTGTGCTGATGC 900
 QY 2154 CGTGTTCAGGCTCATGAGACAGTCTGTGGGCTTATGA 2189
 DB 901 CGTGTTCAGGCTCATGAGACAGTCTGTGGGCTTATGA 936
 RESULT 4
 AK004587
 LOCUS 3373 bp mRNA linear HTC 03-APR-2004
 DEFINITION Mus musculus adult male lung cDNA, RIKEN full-length enriched
 library, clone:1200003015 product:similar to PROTO-ONCOGENE
 TYROSINE-PROTEIN KINASE FES/FPS (BC 2.7.1.112) (C-FES) [Mus
 musculus], full insert sequence.
 AK004587
 ACCESSION
 VERSION AK004587.1 GI:12835858
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multiplexed capillary sequencer
 JOURNAL Nature Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3373)

AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arono,H., Atakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hirooka,T., Horii,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Keskula,T., Kato,H., Kawai,Y., Koijima,Y., Komori,H., Konda,M., Koyama,S., Kuwahara,C., Matsuyama,T., Miyazaki,A., Nishikawa,S., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.jp, URL:<http://genome.gsc.riken.jp/>, Fax:81-45-503-9216)

COMMENT Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAAGAGCGCCGCAACTGCAGCTTTTATTTTTTAA 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAAGAGCATCAGACGTCAATTAAATTAACCCTCCCC 3'. cDNA was cleaved with XhoI and Sall. Cloning sites, 5' end; Sall; 3' end; XhoI. Host: SOLR.

FEATURES

source	location/qualifiers 1..3373 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM.DB:1200003015" /db_xref="taxon:10090" /clone="1200003015" /sex="male" /tissue_type="lung" /clone_libs="RIKEN full-length enriched mouse cDNA library" /dev_abbrev="adult"
misc_feature	1..3373 /note="Similar to PROTO-ONCOGENE TYROSINE-PROTEIN KINASE PES/FPS (EC 2.7.1.112) (C-FES) [Mus musculus] (SWISSPROT P16879, evidence: FASTV, 77.2%ID, 100%length, match=3083)"

ORIGIN

Query Match	39.2%; Score 885.4; DB 3; Length 3373;
Best Local Similarity	77.8%; Pred. No. 5.6e-176;
Matches 1146; Conservative	0; Mismatches 191; Indels 136; Gaps 1;

```

QY      1 ATGGGCTTCCTTGAGCTGTGACGCCCGGAGGCCAAGCGGTCCTGACACAATGAG 60
DB      112 ATGGGCTTCCTTGAGCTGTGACGCCCGGAGGCCAAGCGGTCCTGACACAATGAG 171
QY      61 GAGGCCAGAGCTCGCTTAAGAGAGGCGCATGAAGAAGTAGATGAGCCAGCGGATCAAGAGT 120
DB      172 GAAGCTAGAGCTGCCCTTAATGAGAGGCGCATGAAGAAGTAGATGAGCCAGAGGATCAAGAGT 231
QY      121 GACAGGAGATATGACAGATCTCTTCAACCAATGTCCTGACAGACAGTGGGGGCCAAGGC 180
  
```

D	b	232	GA	CCGGGAATATG	CAAGATTG	CTTACCA	CATGTC	CTTTG	CAAGGA	CACGGAGGCC	CAAGAC	291
Q	y	181	CGGG	CCATCAG	CCCTCGA	CAGCCCAT	CAGTCAGTC	CTGGGCTG	AGATC	CAACGCAAA	CT	240
D	b	292	TG	AGCAG	CCGGCC	CCGACAG	CCCTGTC	AGCAAGTC	CTGGGCA	GATATACA	AGCCAGAC	351
Q	y	241	GAGG	AGCTG	AGCCGCTT	GCTG	CCGGCAG	CAGCAG	AGGATCT	GAACTC	GAGGCCCTTG	300
D	b	352	GAGAA	CTTGA	GC	CCGGGCTG	CCGGCAG	CATGCA	GAATCT	GAACTCGGGCC	CTTGAGC	411
Q	y	301	AAGCTG	AGCCGTC	ATCCCGG	GAACGG	CAGAGCTT	GGCAAGA	CTTCA	AGGCAAGG		360
D	b	412	AA	CTGAGC	CTGTG	ATCTG	ATCCGAG	CCGAGCA	CTTG	GAAGAAGT	CTAC	471
Q	y	361	CAGCAG	CTGAG	CAGAG	AGCTCA	CCCAAG	CCCA	CAGCAG	CATTG	GAAGCTG	420
D	b	472	CAGAG	CTGAG	CAGAG	AGCTCA	CCCAAG	CCCA	CAGCAG	CATTG	GAAGCTG	531
Q	y	421	CAGTA	CCGAG	CTG	GCAC	CGGGA	CAGTC	AGCC	CAAGTC	CAAGAGCC	480
D	b	532	CAGTAC	CGGAC	CCCTG	GTACAG	ATAG	AGCACC	AGGCC	CAGGCAAGT	ACAGGAAGCC	591
Q	y	481	AA	AGACA	AGGAC	CCGTCG	CAAGG	CCCA	AGGA	CAAGTA	TGTCG	540
D	b	592	AA	AAACA	AGAGAT	CCAGAC	AAAG	CCCAAG	CAAGAT	GTCCG	AGCTGTG	651
Q	y	541	GCTC	ACCA	CAAC	CGCTAT	GTGCTG	GGGCGT	GC	CGGCTCG	CACTAC	600
D	b	652	GCCCA	CCACA	CAAC	CGCTAT	GTGCTG	GGGCGT	GC	CGGCTCG	CACTAC	711
Q	y	601	CACCA	GCTC	CTGCTG	CCCGG	CTGCTG	CGGT	CACTG	CAAG	CACTG	660
D	b	712	CACG	GTTCAT	GTGCTG	CCCGG	CTGCTG	CGGT	CACTG	CAAG	CACTG	771
Q	y	661	TG	CACTCT	GAAAG	GAATCT	CGCAG	GAATTA	CTTG	GAATTA	AGAC	720
D	b	772	GGCAT	TTG	GAAGACAT	CTCG	CAGAA	TAATCTG	GAATTA	AGAC	CTG	831
Q	y	721	GTCG	TG	CCATTC	CAAC	CGGAG	ATG	GGCTG	CGAC	CTGCT	780
D	b	832	GTG	CAATCC	ATTC	CAAC	CGGAG	ATG	GGCTG	CGAC	CTGCT	891
Q	y	781	TACCA	GAGCTT	CCTG	CGAC	AGTATG	GGT	CGAC	CTGAC	GTCC	840
D	b	892	TACCT	AGGCTT	CCG	CGAC	ATATG	ATG	ATCA	CCCTG	ATG	951
Q	y	841	GATG	ATC	ATC	ATGCTT	GAGAG	GGTGA	AC	CCCTG	AGG	900
D	b	952	GATG	ATC	ATC	ATGCTT	GAGAG	GGTGA	AC	CCCTG	AGG	1011
Q	y	901	CTG	ACTT	GAGAG	CGG	CAGCA	CAC	-----	-----	-----	926
D	b	1012	TTG	ACTT	GAGAG	CGG	CAGCA	CAC	CGTGT	ACTCT	CTGT	1071
Q	y	927	-----	-----	-----	-----	-----	-----	-----	-----	-----	926
D	b	1072	GTG	GGC	CAGGAG	CCCTT	GGC	CAGG	GTG	GGCTG	GTG	1131
Q	y	927	-----	-----	-----	-----	-----	-----	-----	-----	-----	944
D	b	1132	TG	GTG	GGG	AGCC	CAGCTT	CTC	CTG	CCCTG	GGG	1191
Q	y	945	TGA	AGCTG	GTG	GGC	CAAC	CGAG	ATG	GTG	TTA	1004
D	b	1192	TGA	AGCTG	GTG	GGC	CAAC	CGAG	ATG	GTG	TTA	1251
Q	y	1005	ACA	GA	AGCTT	CCG	GAATGA	AGG	GAAC	CA	CCCA	1064
D	b	1252	GCG	GA	AGCTT	CCG	GAATGA	AGG	GAAC	CA	CCCA	1311
Q	y	1065	CA	AG	AGCA	AGTGTG	TCGA	AG	AG	AG	AG	1124
D	b	1312	CA	AG	AGCA	AGTGTG	TCGA	AG	AG	AG	AG	1371

Oy	1125	CAAGCTGAGGGCCCGACAGAGAGTTGTCGCACAACCAAGTGGAGACCCTGGGCCGCCGGGGA	118
Db	1372	CAACTTCATGAGCCCGACAGAGAACTTCTTGACAGCAGAATGAGACAGCTCGGCACGGGGCA	1431
Oy	1185	GCCCCCGCTGTGTGCTCTCTCTGACAGATGACCGGCACCTCCACGTCGTCTTGGAGCAGA	1244
Db	1432	GCCCCCAGCGCTGC CGCTCTTTCAGAGATGACCGGCATTCCACTCTCTCACGAGCAGA	1491
Oy	1245	GCGAGAGGGGGGAGAGACACCCAGCGCTGAGATCTCTTAAGGCCACCATCTTAGAATCTT	1304
Db	1492	GCGAGAGGGGGGAGAGACACCCACCACTGGAATCTTTAAAAGCCACTTCTCCGAATCTT	1551
Oy	1305	CCGCCCCCAAGTTCTCGAACCTGTACCCGACTGGA	1337
Db	1552	CCGCCCCCAAGTTCTCTATCTTATCTCCACCACTGCA	1584
RESULT 5			
AL542976		936 bp	mRNA linear EST 24-MAR-2004
LOCUS			
DEFINITION	AL542976 Homo sapiens PLACENTA Homo sapiens cDNA clone CSDBE013YP02		
ACCESSION	5-PRIME, mRNA sequence.		
VERSION	AL542976		
KEYWORDS	AL542976.3 GI:45718543		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	On Feb 15, 2001 this sequence version replaced gi:30548660. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gascon Cremieux, CP 5706 - 91057 Evry cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7663.f For more information about this cluster, see http://www.genoscope.cns.fr/cdna76=CSDBE013DHO1QP1&c=7663.f .		
FEATURES			
source	Location/Qualifiers		
	1..936		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CSDBE013YP02"		
	/tissue_type="PLACENTA"		
	/clone_lib="Homo sapiens PLACENTA"		
	/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."		
ORIGIN			
Query Match	37.6%; Score 848.8; DB 1; Length 936;		
Best Local Similarity	99.2%; Pred. No.2,5e-168;		
Matches	861; Conservative 2; Mismatches 4; Indels 1; Gaps 1;		
Oy	1 ATGGGCTTCTCTTGAAGCTGTGACGCCCCCAGGCGCACGAGGCTCTTGACCAATGTCAG	60	
Db	57 ATGGGCTTCTCTTGAAGCTGTGACGCCCCCAGGCGCACGAGGCTCTTGACCAATGTCAG	116	
Oy	61 GAGGCGGAGCTTCTGCTCTACTGAGAGGCGCATGAGAAAGTGGATGGCCAGCGGGTCAAAGT	120	
Db	117 GAGGCGGAGCTTCTGCTCTACTGAGAGGCGCATGAGAAAGTGGATGGCCAGCGGGTCAAAGT	176	

Oy	121	GACAGGGAAGTATGAGGAGTCTTCAACAATGTCCTGACAGGACATGGGGCCAGAGC	180
Db	177	GACAGGGAAGTATGAGGAGTCTTCAACAATGTCCTGACAGGACATGGGGGCCAAGC	236
Oy	181	CGGGCCATCAGCCCTGACAGCCCCATCATGTCATCTTGGGCTGAGATCAACGCAAACT	240
Db	237	CGGGCCATCAGCCCTGACAGCCCCATCATGTCATCTTGGGCTGAGATCAACGCAAACT	296
Oy	241	GAGGGCTTGAACCGCTTGGCTGGCGGACAGCAAGGATCTGAATCTAGGGCCCTGAGC	300
Db	297	GAGGGCTTGAACCGCTTGGCTGGCGGACAGCAAGGATCTGAATCTAGGGCCCTGAGC	356
Oy	301	AAGGTGAGCCTGTGTCATCCGGGAAACGGCAGCAGTTTGGCAAGACCTTACAGGAGCAGTGG	360
Db	357	AAGGTGAGCCTGTGTCATCCGGGAAACGGCAGCAGTTTGGCAAGACCTTACAGGAGCAGTGG	416
Oy	361	CAGCAGCTGACAGCAGAGCTTCAACAGACCCACAGCCAGCATTTGAGAACTGAGAGC	420
Db	417	CAGCAGCTGACAGCAGAGCTTCAACAGACCCACAGCCAGCATTTGAGAACTGAGAGC	476
Oy	421	CAGTACCGAGCTCTGGGACCGGGAACAGTCCCAAGGCCAAGGCCAAGTACAGAGAGCCAGC	480
Db	477	CAGTACCGAGCTCTGGGACCGGGAACAGTCCCAAGGCCAAGGCCAAGTACAGAGAGCCAGC	536
Oy	481	AAAGACAAGGACCTGTACAAGGCCCAAGAGACAATGATGTGGCAGAGCTGTGAACTCTTT	540
Db	537	AAAGACAAGGACCTGTACAAGGCCCAAGAGACAATGATGTGGCA-SCTGTGAACTCTTT	595
Oy	541	GCTCACAACAACCGCTATGTGTGTGGGCGTGGCGGCTGGCCAGCTTACACACACACAC	600
Db	596	GCTCACAACAACCGCTATGTGTGTGGGCGTGGCGGCTGGCCAGCTTACACACACACAC	655
Oy	601	CACCAAGCTCTGCTGCGCCCGGCTGTGCGGATCATGTGAGAGACCTGACAGAGATGGCT	660
Db	656	CACCAAGCTCTGCTGCGCCCGGCTGTGCGGATCATGTGAGAGACCTGACAGAGATGGCT	715
Oy	661	TGCATCTCTGAAGAGATCTTGCAGAAATACCTTGAGATTAGCAGCCTGTGTACAGATGAG	720
Db	716	TGCATCTCTGAAGAGATCTTGCAGAAATACCTTGAGATTAGCAGCCTGTGTACAGATGAG	775
Oy	721	GTGTGTGTCATTTCAACCGGAGATATGGCTGTGCACTGTCTGCCCATTCAGAGCTGAG	780
Db	776	GTGTGTGTCATTTCAACCGGAGATATGGCTGTGCACTGTCTGCCCATTCAGAGCTGAG	835
Oy	781	TACCAAGGCTTCCATCGACACAGTATGGGTGCGCAGCTTACAGCTTGTGTACGTTT	840
Db	836	TACCAAGGCTTCCATCGACACAGTATGGGTGCGCAGCTTACAGCTTGTGTACGTTT	895
Oy	841	GATGAGTCACTGCTTGAAGAGGATGAAC	868
Db	896	GATGAGTCACTGCTTGAAGAGGATGAAC	923
RESULT 6			
LOCUS	BM456755	1029 bp	mRNA linear EST 05-FEB-2002
DEFINITION	AGENCOURT 6403957 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583652		
ACCESSION	BM456755		
VERSION	BM456755.1	GI:18505795	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: gsapbs@remail.nih.gov		
	Tissue Procurement: ATCC		

cdna library Preparation: Life Technologies, Inc.
 cdna library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM12347 row: b column: 05
 High quality sequence stop: 704.

FEATURES

source

Location/Qualifiers

1..1029

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5583552"

/issue_type="embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_92"

/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."

ORIGIN

Query Match 35.7%; Score 805.8; DB 4; Length 1029;
 Best Local Similarity 97.9%; Pred. No. 3e-159; Indels 6; Gaps 5;
 Matches 869; Conservative 0; Mismatches 13;

QY 1 ATGGGCTTCTCTCTGAGCTGTGACAGCCGCCAGGCGGAGGCTCTGACGAAATGACG 60
 DB 81 ATGGGCTTCTCTCTGAGCTGTGACAGCCGCCAGGCGGAGGCTCTGACGAAATGACG 140
 QY 61 GAGGCCGAGCTTCTCTGAGCTGTGAGGCGGAGGAGGATGAGTGGCCGAGGCTCAAGAGT 120
 DB 141 GAGGCCGAGCTTCTCTGAGCTGTGAGGCGGAGGAGGATGAGTGGCCGAGGCTCAAGAGT 200
 QY 121 GACAGGAGATGATGACAGGAGCTTCAACATGCTCCCTGAGGACAGTGCGGCGGCGAGAGC 180
 DB 201 GACAGGAGATGATGACAGGAGCTTCAACATGCTCCCTGAGGAGAGTGCGGCGGCGAGAGC 260
 QY 181 CGGGCCATGACCCCTGACAGCCCATCATGCTCTGAGGCTGAGATCAACAGCAAACT 240
 DB 261 CGGGCCATGACCCCTGACAGCCCATCATGCTCTGAGGCTGAGATCAACAGCAAACT 320
 QY 241 GAGGGCTTGAAGCGCTTGTCTGCGGACAGCAGCAGAGAGATGTAATCAGGCGCCCTGAGC 300
 DB 321 GAGGGCTTGAAGCGCTTGTCTGCGGACAGCAGCAGAGAGATGTAATCAGGCGCCCTGAGC 380
 QY 301 AAGCTGAGCTGCTCTTCCGGGAAACGGGACAGCTTCCGAAACCTTACAGCCGAGCAGTGG 360
 DB 381 AAGCTGAGCTGCTCTTCCGGGAAACGGGACAGCTTCCGAAACCTTACAGCCGAGCAGTGG 440
 QY 361 GACGAGCTCAGAGAGAGCTCAACAGAGCCAGCAGCAGGACATTTGAGAACTGAGAGC 420
 DB 441 GACGAGCTCAGAGAGAGCTCAACAGAGCCAGCAGCAGGACATTTGAGAACTGAGAGC 500
 QY 421 CAGTACCTGAGCTCTGCGACGGGACAGTGCCCAAGCCAAAGGCAAGTACCGAGAGCCAGC 480
 DB 501 CAGTACCTGAGCTCTGCGACGGGACAGTGCCCAAGCCAAAGGCAAGTACCGAGAGCCAGC 560
 QY 481 AAAGAAAGAGCCGTGACAAAGGCAAGGAACTAATGTCGGGAGCCCTGTGGAAGCTTTT 540
 DB 561 AAAGAAAGAGCCGTGACAAAGGCAAGGAACTAATGTCGGGAGCCCTGTGGAAGCTTTT 620
 QY 541 GCTCACCACAAACGCTATGCTGAGGCGGTGCGGAGCTGACAGCAACAGCAGCAGC 600
 DB 621 GCTCACCACAAACGCTATGCTGAGGCGGTGCGGAGCTGACAGCAACAGCAGCAGC 680
 QY 601 CACCAAGCTCTGCTGCGGCGCTGCTGCGGTGCTGACAGGACCTGACAGAGAGTGGCT 660
 DB 681 CACCAAGCTCTGCTGCGGCGCTGCTGCGGTGCTGACAGGACCTGACAGAGAGTGGCT 740
 QY 661 TGCATCTGAAGAGATCTCTGCAAGAAATACCTGGAATTAAGCAGCCTGCTGCAAGAGTGA 720

DB 741 TGCATCTGAAGAGATCTCTGCAAGAAATACCTGGAATTAAGCAGCCTGCTGCAAGAGTGA 800
 QY 721 GGTGGGCGCATTCACCGGAGG-AGATGGCTGACAGCTGTGCGCGCATTCAGCTGAGCTGA 779
 DB 801 GGTGGGCGCATTCACCGGAGGAGTGGCTGACAGCTGTGCGCGCATTCAGCTGAGCTGA 860
 QY 780 GTACCAAGGCTTCTGCGACAGTATGAGTCCGACCTGAC-GTCCCAACCTGTGTACG- 837
 DB 861 GTACCAAGGCTT-CTGCGACAGTATGAGTCCGACCTGACAGCTGCAACCTGTGTACG 919
 QY 838 TTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 883
 DB 920 TTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 967

RESULT 7
 BX359010 968 bp mRNA linear EST 08-APR-2004
 LOCUS BX359010 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1051Y020 5-PRIME, mRNA sequence.
 ACCESSION BX359010
 VERSION BX359010.2 GI:46307488
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polyes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On May 5, 2003 this sequence version replaced gi:30378289.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7663.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?b=CS0D1051BH10P1&c=7663.f>.
 Location/Qualifiers
 1..968

FEATURES

source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1051Y020"

/issue_type="PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 34.7%; Score 783.6; DB 5; Length 968;
 Best Local Similarity 95.9%; Pred. No. 1.4e-154; Indels 9; Gaps 8;
 Matches 879; Conservative 9; Mismatches 20;

QY 1 ATGGGCTTCTCTCTGAGCTGTGACAGCCGCCAGGCGGAGGCTCTGACGAAATGACG 60
 DB 61 ATGGGCTTCTCTCTGAGCTGTGACAGCCGCCAGGCGGAGGCTCTGACGAAATGACG 120
 QY 61 GAGGCCGAGCTTCTCTGAGCTGTGAGGCGGAGGAGGATGAGTGGCCGAGGCTCAAGAGT 120
 DB 121 GAGGCCGAGCTTCTCTGAGCTGTGAGGCGGAGGAGGATGAGTGGCCGAGGCTCAAGAGT 180
 QY 121 GACAGGAGATGATGACAGGAGCTTCAACATGCTCCCTGAGGACAGTGCGGCGGCGAGAGC 180

```

Db      181 GACGAGGAGTATGACAGACTGCTTCAACAATGTCCTGTGACGAGACAGTGGGGCCAGAGC 240
QY      181 |CGGGCCATCATGAGCCCTGACAGCCCCATCAAGTCACTCCGCGCTGAGATCAACAGCCAAACT 240
Db      241 CGGGCCATCATGAGCCCTGACAGCCCCATCAAGTCACTCCGCGCTGAGATCAACAGCCAAACT 300
QY      241 GAGGGCCTGAGCCGCTTGTGCTGGGCGACAGCAGAGAGATCTGAATCTCAGGGCCCTGAGC 300
Db      301 GAGGGCCTGAGCCGCTTGTGCTGGGCGACAGCAGAGAGATCTGAATCTCAGGGCCCTGAGC 360
QY      301 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACAGCTTGGCAAGCTTAACAGCCAGCAGTGG 360
Db      361 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACAGCTTGGCAAGCTTAACAGCCAGCAGTGG 420
QY      361 CAGCAGCTGACAGAGAGAGCTCACCAAGACCCACAGCAGAGACATTTGAGAGCTGAAGAGC 420
Db      421 CAGCAGCTGACAGAGAGAGCTCACCAAGACCCACAGCAGAGACATTTGAGAGAGCTGAAGAGC 480
QY      421 CAGTACCGAGCTTGGCAGCGGGAAGTCCCAAGCCCAAGCCCAAGAGTACAGAGAGCCAGC 480
Db      481 CAGTCCGAGCTTGGCAGCGGGAAGTCCCAAGCCCAAGAGTACAGAGAGCCAGC 540
QY      481 AAGACACAGGACCGCTGACAGAGCCCAAGAGACAAATATGTGCGCAGCTGTGAAAGCTCTTT 540
Db      541 AAGACACAGGACCGCTGACAGAGCCCAAGAGACAAATATGTGCGCAGCTGTGAAAGCTCTTT 599
QY      541 GCTCACACACAGCGCTATGTGCTGGGCGTGGGGCTGGCCAGCTACACCAACAGCAGCCAC 600
Db      600 GCTCACACACAGCGCTATGTGCTGGGCGTGGGGCTGGCCAGCTACACCAACAGCAGCCAC 657
QY      601 CACACAGCTCTGCTGGCCCGGCTGCTGCGGTCACTGACAGAGCCTGACAGAGAGATGGCT 660
Db      658 CACACAGCTCTGCTGGCCCGGCTGCTGCGGTCACTGACAGAGCCTGACAGAGAGATGGCT 717
QY      661 TGCACTCTGAAGAGATCTCTGACAGAAATACCTGGAATTTAGCAGCCTGTGTGACAGATGAG 720
Db      718 TGCACTCTGAAGAGATCTCTGACAGAAATACCTGGAATTTAGCAGCCTGTGTGACAGATGAG 777
QY      721 GTGGTGGCCATTACACGGGAGATGAGGCTGACAGCTGCTGCCCGCATCCAGCCTGAGGCTGAG 780
Db      778 GTGGTGGCCATTACACGGGAGATGAGGCTGACAGCTGCTGCCCGCATCCAGCCTGAGGCTGAG 834
QY      781 TACCAAGGCTTCTGCGGACAGTATGGGTCCGACCTGACCTCCACCTGTGTCAAGTTTC 840
Db      835 TA-CAAGGCTTCTGCGGACAGTATGGGT-CCGACCTGACCTCCACCTGTGTCAAGTTTC 892
QY      841 GATGAGTCACTGCTTGAAGAGGGTGAACCGCTGGAAGCTTCCAGCTGAACGAG 900
Db      893 GATGAGTCACTGCTTGAAGAGGGTGAACCGCTGGAAGCTTCCAGCTGAACGAG 951
QY      901 CTGACTGTGAGAGCGT 917
Db      952 CTGACTGTGAGAGAGT 968

```

```

RESULT 8
BX381726 939 bp mRNA linear EST 26-Apr-2004
LOCUS BX381726 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS001069YK21 5-PRIME, mRNA sequence.
ACCESSION BX381726
VERSION BX381726.2 GI:46570766
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 939)
AUTHORS Li, W.B., Gruber, C., Jessup, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30443719.
Contact: Genoscope

```

Genoscope - Centre National de Séquençage
 2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
 Email: secref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with NotI and cloned
 into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by life technologies, a
 division of invitrogen. This sequence belongs to sequence cluster
 7663.f
 For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?as=CS001069AF110P1ac=7663.f>.
 Location/Qualifiers

FEATURES

```

source
1..939
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS001069YK21"
/cisue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match 34.3%; Score 773.4; DB 5; Length 939;
 Best local Similarity 97.2%; Pred. No. 2e-152;
 Matches 845; Conservative 13; Mismatches 4; Indels 7; Gaps 7;

```

QY      1 ATGGGCTTCTCTTCTGAGCTGTGACAGCCCGCCAGCGGCTCTGACCAATATGACG 60
Db      32 ATGGGCTTCTCTTCTGAGCTGTGACAGCCCGCCAGCGGCTCTGACCAATATGACG 91
QY      61 GAGGCGGAGCTTCTGCTGAGAGGCGATGAGAAATGATGGCCAGCGGGCTCAAGAT 120
Db      92 GAGGCGGAGCTTCTGCTGAGAGGCGATGAGAAATGATGGCCAGCGGGCTCAAGAT 151
QY      121 GACAGGAGTATGACAGACTGCTTCAACATATGCTCCCTGACAGAGCAGTGGGGCCAGAGC 180
Db      152 GACAGGAGTATGACAGACTGCTTCAACATATGCTCCCTGACAGAGCAGTGGGGCCAGAGC 211
QY      181 CGGGCCATCATGAGCCCTGACAGCCCCATCAAGTCACTCCGCGCTGAGATCAACAGCCAACT 240
Db      212 CGGGCCATCATGAGCCCTGACAGCCCCATCAAGTCACTCCGCGCTGAGATCAACAGCCAACT 271
QY      241 GAGGGCCTGAGCCGCTTGTGCTGGGCGACAGCAGAGAGATCTGAATCTCAGGGCCCTGAGC 300
Db      272 GAGGGCCTGAGCCGCTTGTGCTGGGCGACAGCAGAGAGATCTGAATCTCAGGGCCCTGAGC 331
QY      301 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACAGCTTGGCAAGCTTAACAGCAGAGTGG 360
Db      332 AAGCTGAGCTGCTCATCCGGGAAACGGCAGACAGCTTGGCAAGCTTAACAGCAGAGTGG 391
QY      361 CAGCAGCTGACAGAGAGAGCTCACCAAGACCCACAGCAGAGACATTTGAGAGCTGAAGAGC 420
Db      392 CAGCAGCTGACAGAGAGAGCTCACCAAGACCCACAGCAGAGACATTTGAGAGCTGAAGAGC 451
QY      421 CAGTACCGAGCTTGGCAGCGGGAAGTCCCAAGCCCAAGAGGCAAGTACCAAGAGGCGCAGC 480
Db      452 CAGTACCGAGCTTGGCAGCGGGAAGTCCCAAGCCCAAGAGGCAAGTACCAAGAGGCGCAGC 511
QY      481 AAGACACAGGACCGCTGACAGAGCCCAAGAGACAAATATGTGCGCAGCTGTGAAAGCTCTTT 540
Db      512 AAGACACAGGACCGCTGACAGAGCCCAAGAGACAAATATGTGCGCAGCTGTGAAAGCTCTTT 570
QY      541 GCTCACACACAGCGCTATGTGCTGGGCGTGGGGCTGGCCAGCTTAACCAACAGCAGCCAC 600
Db      571 GCTCACACACAGCGCTATGTGCTGGGCGTGGGGCTGGCCAGCTTAACCAACAGCAGCCAC 629
QY      601 -CACGAGCTCTGCTGGCCCGGCTGTGGGCTCACTGACAGAGCTTCAAGAGAGATGGC 659
Db      630 AACACGAGCTCTGCTGGCCCGGCTGTGGGCTCACTGACAGAGCTTCAAGAGAGATGGC 689

```

QY 660 TTGCATCTGAAGAGATCCTGCGAGAAATCCTGAGATTAGAGAGCTGGTGCAGATGA 719
DB 690 TTGCATCTGAAGAGATCCTGCGAGAAATCCTGAGATTAGAGAGCTGGTGCAGATGA 749
QY 720 GGTGTGGCCATTCAACCGGAGATGCTGCACTGCTGCCGATCCAGCTTGAAGCTGA 779
DB 750 GGTGTGGCCATTCAACCGGAGATGCTGCACTGCTGCCGATCCAGCTTGAAGCTGA 809
QY 780 GATCAAGGCTTCTGCGACAGTATGGGTCGCCACTGACGTCCACCTGTGTACGTT 839
DB 810 GATCAAGGCTTCTGCGACAGTATGGGTCGCCACTGACGTCCACCTGTGTACGTT 866
QY 840 CGATGATCACTGCTGAGAGAGGTGAAC 868
DB 867 CGATG-RKTMATGCTTKAAGAGGTGAAC 894

RESULT 9
B0708270 948 bp mRNA linear EST 16-JUL-2002
LOCUS AGENCOURT_8484911 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6301278
DEFINITION 5', mRNA sequence.
ACCESSION B0708270
VERSION B0708270.1 GI:21847169
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 948)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2516 row: 0 column: 07
High quality sequence stop: 610.
Location/Qualifiers
1. 948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="6301278"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_113"
/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGGACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 34.2%; Score 771.6; DB 5; Length 948;
Best Local Similarity 95.3%; Pred. No. 4.8e-152;
Matches 818; Conservative 0; Mismatches 34; Indels 6; Gaps 2;
QY 223 GAGATACCAAGCAAACTGAGGGGCTGAGCCGCTTCTGCGGAGAGACGAGAGATTCG 282
DB 1 GAGATACCAAGCAAACTGAGGGGCTGAGCCGCTTCTGCGGAGAGACGAGAGATTCG 60
QY 283 AACTCAGGGCCCTGAGCAAGTGAAGCTGCTCATCCGGGAACGAGCAGCTTCCCAAG 342
DB 61 AACTCAGGGCCCTGAGCAAGTGAAGCTGCTCATCCGGGAACGAGCAGCTTCCCAAG 120

QY 343 ACTTACAGGAGAGTGGAGAGAGCTGACAGAGAGAGTCAACCAAGACCAAGCCAGAG 402
DB 121 ACTTACAGGAGAGTGGAGAGAGCTGACAGAGAGAGTCAACCAAGACCAAGCCAGAG 180
QY 403 ATTGAGAGCTGAAGAGAGTGAACCGAGCTTGGACAGGAGACAGTCCCAAGCAAGCG 462
DB 181 ATTGAGAGCTGAAGAGAGTGAACCGAGCTTGGACAGGAGAGTCCCAAGCAAGCG 240
QY 463 AAGTACAGAGGCGCAGCAAAAGCAAGACCGTGAACAAGCCCAAGCAAGATATGCGC 522
DB 241 AAGTACAGAGGCGCAGCAAAAGCAAGACCGTGAACAAGCCCAAGCAAGATATGCGC 300
QY 523 AGCTGTGAGAGTCTTTGTCTCAACCAACCGTATATGCTGGGCGTGGGCTGGGAG 582
DB 301 AGCTGTGAGAGTCTTTGTCTCAACCAACCGTATATGCTGGGCGTGGGCTGGGAG 360
QY 583 CTACACCAAGAGCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 642
DB 361 CTACACCAAGAGCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 643 CTGACAGAGAGATGAGCTTGCATCTGAAAGAGATCTGCAAGAAATACCTGAGATTAGC 702
DB 421 CTGACAGAGAGATGAGCTTGCATCTGAAAGAGATCTGCAAGAAATACCTGAGATTAGC 480
QY 703 AGCTGTGAGAGATGAGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 762
DB 481 AGCTGTGAGAGATGAGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 763 ATCCAGCTGAGGCTGAGTACCAAGGCTTCTGCGCAGATGAGTCCGACCTGACGTC 822
DB 541 ATCCAGCTGAGGCTGAGTACCAAGGCTTCTGCGCAGATGAGTCCGACCTGACGTC 600
QY 823 CCAACCTGTGTCACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 882
DB 601 CCAACCTGTGTCACTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
QY 883 GAGCTCAGCTGAACAGAGCTGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTG 942
DB 661 GAGCTCAGCTGAACAGAGCTGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTG 720
QY 943 GATGAGCTGAGTGTGAGAGAGTGTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1000
DB 721 GATGAGCTGAGTGTGAGAGAGTGTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 780
QY 1001 TGCAACAGAGAGCTCCGAAATGAAGAGA----GAACACCAACCCCGGAGACGGGTG 1056
DB 781 TGCAACAGAGAGAGCTCGAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 1057 CTGCTGGGCAAGAGGCA 1074
DB 841 CAGCCCTGCTGGGGCCAA 858

RESULT 10
BX356088/c 980 bp mRNA linear EST 08-APR-2004
LOCUS BX356088 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1008YG07 3-PRIME, mRNA sequence.
ACCESSION BX356088
VERSION BX356088.2 GI:46288403
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 980)
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30369992.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage

QY 560 TGGTGGGCGTGGGGGCTGGGCGAGCTACAGACCAACAGACCAACCACTCTGTGGCG 619
 Db 249 TGGTGGGCGTGGGGGCTGGGCGAGCTACAGACCAACAGACCAACCACTCTGTGGCG 308
 QY 620 GCGTGTGCGGTGCTGCTGAGAGCTGAGAGAGATGAGTGTGATCTCTGATAGAGATCC 679
 Db 309 GCGTGTGCGGTGCTGCTGAGAGCTGAGAGAGATGAGTGTGATCTCTGATAGAGATCC 368
 QY 680 TGCAGAGATACCTGAGAGATTAGAGAGCTGTGTGAGAGATGAGTGTGAGAGATCC 739
 Db 369 TGCAGAGATACCTGAGAGATTAGAGAGCTGTGTGAGAGATGAGTGTGAGAGATCC 428
 QY 740 AGATGTGTGAGAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 799
 Db 429 AGATGTGTGAGAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 488
 QY 800 AGATGTGTGAGAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 859
 Db 489 AGATGTGTGAGAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 548
 QY 860 AGGATGTGTGAGAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 919
 Db 549 AGGATGTGTGAGAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 608
 QY 920 AGCAGAGCTGAGCTGCTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 979
 Db 609 AGCAGAGCTGAGCTGCTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 668
 QY 980 GGCAGAGATGAGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 1037
 Db 669 GGCAGAGATGAGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 728
 QY 1038 CCCCCTGGAGAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 1097
 Db 729 CCCCCTGGAGAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 788
 QY 1098 GCTGAGAGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 1155
 Db 789 GCTGAGAGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 848
 QY 1156 ACCAAGCT-GGAGAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 1210
 Db 849 AACCAAGCTGTGTGAGAGATGAGTGTGAGAGATGAGTGTGAGAGATCC 908
 QY 1211 ATGACCGGCACTC 1223
 Db 909 AAGAACCCCCAC 921

RESULT 12
 BMS43870 1114 bp mRNA linear EST 20-FEB-2002
 LOCUS AGENCOURT 6492411 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5589213
 DEFINITION 5' mRNA sequence.
 ACCESSION BMS43870.1 GI:18774639
 VERSION BMS43870.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1114)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsdbs-f@mail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LIML at:
<http://image.jnl.gov>
 Plate: LHAM12361 Row: 1 Column: 22
 High quality sequence stop: 566.
 Location/Qualifiers
 1. 1114
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5589213"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_125"
 /note="Organ: ovary (pool of 3); Vector: pCMV-Sport6;
 Site 1: Ecorev (destroyed); Site 2: NotI; RNA source pool
 of three ovaries, from females ranging in age from 38 to
 49 yo. Library is oligo-dT primed and directionally cloned
 (Ecorev site is destroyed upon cloning). Average insert
 size 2.1 kb, insert size range 1-3.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 036."

ORIGIN

Query Match 31.5%; Score 710.6; DB 4; Length 1114;
 Best Local Similarity 95.7%; Pred. No. 3.6e-139;
 Matches 794; Conservative 0; Mismatches 27; Indels 9; Gaps 6;

QY 147 CCACATGTCCTGTCAGAGAGAGTGGGGCCAGAGCCGGCCATCAGCCCTGACAGCCCAT 206
 Db 12 CCACATGTCCTGTCAGAGAGAGTGGGGCCAGAGCCGGCCATCAGCCCTGACAGCCCAT 71
 QY 207 CAGTCACTCTGCGGTGAGATCAACAGCCAACTGAGGCTGAGCCGCTTGTGCGCA 266
 Db 72 CAGTCACTCTGCGGTGAGATCAACAGCCAACTGAGGCTGAGCCGCTTGTGCGCA 121
 QY 267 GCAAGAGAGATCTGAACTCAGAGGCGCTGAGCAAGCTGAGCTGCTCATCCGGGAAG 326
 Db 132 GCAAGAGAGATCTGAACTCAGAGGCGCTGAGCAAGCTGAGCTGCTCATCCGGGAAG 191
 QY 327 GCAAGAGATCTGAGAGATCTGAGAGAGTGGGAGAGCTGAGAGAGTGCACCA 386
 Db 192 GCAAGAGATCTGAGAGATCTGAGAGAGTGGGAGAGCTGAGAGAGTGCACCA 251
 QY 387 GACCCAGAGAGATCTGAGAGAGTGGGAGAGCTGAGAGAGTGCACCA 446
 Db 252 GACCCAGAGAGATCTGAGAGAGTGGGAGAGCTGAGAGAGTGCACCA 311
 QY 447 TGCCCAAGCAGAGAGATCTGAGAGAGTGGGAGAGCTGAGAGAGTGCACCA 506
 Db 312 TGCCCAAGCAGAGAGATCTGAGAGAGTGGGAGAGCTGAGAGAGTGCACCA 371
 QY 507 GCAAGAGATCTGAGAGAGTGGGAGAGCTGAGAGAGTGGGAGAGTGCACCA 566
 Db 372 GCAAGAGATCTGAGAGAGTGGGAGAGCTGAGAGAGTGGGAGAGTGCACCA 421
 QY 567 CGTGCGGGCTGCGAGAGTCAACCAAGAGAGTGGGAGAGTGGGAGAGTGCACCA 626
 Db 432 CGTGCGGGCTGCGAGAGTCAACCAAGAGAGTGGGAGAGTGGGAGAGTGCACCA 491
 QY 627 GCGGTCACTGAGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGCACCA 686
 Db 492 GCGGTCACTGAGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGCACCA 551
 QY 687 ATACCTGAGATTTAGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGCACCA 746
 Db 552 ATACCTGAGATTTAGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGCACCA 611
 QY 747 TGCAGCTGCTGCGCGATCCAGCTGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGCACCA 806
 Db 612 TGCAGCTGCTGCGCGATCCAGCTGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGCACCA 671
 QY 807 GTCGAGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGGGAGAGTGCACCA 864

Db 672 GTCCGACCTGACCTCCACCTCTGTGTCATGATGACCACTGCTTGAGGAGATT 731
 QY 865 GAACCCCTGGAGCTT-GGGAGCTCCAGC--TGAAGAGCTGACTGTG-AGAGCGTGA 920
 Db 732 AACCCCTGGAGCTGGGGAGCTCCAGCTTGAACCACTGACTGTGGAAGCGGGCA 791
 QY 921 GCACAGC--TGACCTCAGTGAAGATGAGC-TGGCTGTGGCCACGAGA 967
 Db 792 CCACAGCCCTGACCTTATGATGACATGAGCTTGGCTGTGGGACCGAAA 841

RESULT 13
 BU184736 847 bp mRNA linear EST 04-SEP-2002
 LOCUS BU184736
 DEFINITION AGENCOURT_8071838 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6088554
 5', mRNA sequence.
 ACCESSION BU184736
 VERSION BU184736.1 GI:22698720
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 JOURNAL NIH-MGC http://mgs.nci.nih.gov/
 COMMENT 1 (bases 1 to 847)
 Unpublished (1999)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@ds-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LICM2326 row: 0 column: 19
 High quality sequence stop: 721.
 Location/Qualifiers

FEATURES
 source
 1..847
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6088554"
 /issue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resilient)"
 /clone_id="NIH_MGC_112"
 /note="Organ: skin; Vector: pOTR7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAAGCAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC library."

ORIGIN

Query Match 31.2%; Score 703.2; DB 5; Length 847;
 Best Local Similarity 99.4%; Pred. No. 1.3e-137;
 Matches 716; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
 QY 380 TCACCAAGACCCACAGCAGCAGCAGTGAAGCTGAGAGCCAGTACCGAGCTGTGGCAC 439
 Db 69 TCAGTCAGACCCACAGCAGCAGCAGTGAAGCTGAGAGCCAGTACCGAGCTGTGGCAC 128
 QY 440 GGGAGAGTGGCCAGCAGCAGCAGCAGTGAAGCTGAGAGCCAGCAGCAGCAGCAGTGA 499
 Db 129 GGGAGAGTGGCCAGCAGCAGCAGCAGTGAAGCTGAGAGCCAGCAGCAGCAGCAGTGA 188
 QY 500 AGGCGAAGACAGTATGTGGCAGCTGTGGAGAGCTTTTGTCTACACCAACCGCGTATG 559
 Db 189 AGGCGAAGACAGTATGTGGCAGCTGTGGAGAGCTTTTGTCTACACCAACCGCGTATG 248

QY 560 TCCTGGCGCTGGCGGCTGGCGAGCTACACACACACACACACAGCTCTGTGCCCCG 619
 Db 249 TCCTGGCGCTGGCGGCTGGCGAGCTACACACACACACACACAGCTCTGTGCCCCG 308
 QY 620 GCTCTGTGGGCTACTGACAGACCTGACAGAGAGATGGCTTGCATCTGGAAGAGATCC 679
 Db 309 GCTCTGTGGGCTACTGACAGACCTGACAGAGAGATGGCTTGCATCTGGAAGAGATCC 368
 QY 680 TGCAGGAATACCTGGAGATTAACAGCTGTGTCAGAGATGAGGTGGGCTATTCACCGG 739
 Db 369 TGCAGGAATACCTGGAGATTAACAGCTGTGTCAGAGATGAGGTGGGCTATTCACCGG 428
 QY 740 AGATGGCTGACAGCTGTGCTGCCCATTCACAGCTGAGGCTGAGTACCAAGGCTTCTGCGAC 799
 Db 429 AGATGGCTGACAGCTGTGCTGCCCATTCACAGCTGAGGCTGAGTACCAAGGCTTCTGCGAC 488
 QY 800 AGTATGGTCCGACCTGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 859
 Db 489 AGTATGGTCCGACCTGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 548
 QY 860 AGGATGAACCGCTGGAGCCTGGGGAGCTGCTGACCTGAAAGAGCTGAGAGCGTGC 919
 Db 549 AGGATGAACCGCTGGAGCCTGGGGAGCTGCTGACCTGAAAGAGCTGAGAGCGTGC 608
 QY 920 ACACACAGCTGACCTCAGTGAAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 979
 Db 609 ACACACAGCTGACCTCAGTGAAGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 668
 QY 980 GCGAGAGATGTTGTTACGACCTGCTGCAACAGAGCTCGGAAATGAAGAGAAACACCCAC 1039
 Db 669 GCGAGAGATGTTGTTACGACCTGCTGCAACAGAGCTCGGAAATGAAGAGAAACACCCAC 728
 QY 1040 CCC-GGGAGCGGGTGGAGCTGTGGGCAAGAGGCAAGTCTGCAAGAACACTGACGGGG 1098
 Db 729 CCGGGAGCGGGTGGAGCTGTGGGCAAGAGGCAAGTCTGCAAGAACACTGACGGGG 788

RESULT 14
 BX359009/1
 LOCUS BX359009 989 bp mRNA linear EST 08-APR-2004
 DEFINITION BX359009 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1051Y020 3-PRIME, mRNA sequence.
 ACCESSION BX359009
 VERSION BX359009.2 GI:46305682
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 JOURNAL Li, W.B., Gruber, C., Jessup, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 5, 2003 this sequence version replaced gi:30376296.
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7663.f

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0D1051B10NPlc=7663.f.
 Location/Qualifiers

FEATURES
 source
 1..989
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1051Y020"
 /issue_type="PLACENTA COT 25-NORMALIZED"

```

/clone_lib="Homo sapiens PLACEINTRA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	30.9%	Score 697.6;	DB 5;	Length 989;
Best Local Similarity	95.7%	Pred. No. 2e-136;		
Matches 787; Conservative	13;	Mismatches 13;	Indels 9;	Gaps 8

1439 CCAAGCAAGTGGGTCTGAACCATGAGACTGCTGTTGGGTGACGAGATTGACACGGG 1438

989 CCAAGAMAGTGGT--TRACATKAAGCTGTG-TGGGTGAGMARWTT-GACGG 934

QY	1499	GGAACCTTTGGCGAAGTGTTCAGCGGACCGCTCGAGCGGACCAACAC-C	-CTGGTGGCGGGT	155
Db	933	GGAACTTTTGGGAAGTGTTCATGGGACCGCTCGAGCGGACCAACACMC	TGGTGGCGGGT	874
QY	1558	AAGTCTTGTGAGAGACGCTCCCACTGA	CCTCAAAGCCAAATTTCACAGAACGAGG	1611
Db	873	AAGTCTTGTGAGAGACGCTCCCACTGA	CCTCAAAGCCAAATTTCACAGAACGAGG	814
QY	1618	ATCTGAAGCAGTACAGCCACCCCAACATCGCGGCTCTCATTTGGTGTGCA	CCCAAG	1677
Db	813	ATCTGAAGCAGTACAGCCACCCCAACATCGCGCTCTCATTTGGTGTGCA	CCCAAG	754
QY	1678	CAGGCC-ATCTACATCTGATGAGCTTGTGACGGGGGGGCACTTCTGA	CCTTCCTCCG	173
Db	753	CAGCCCAATCTACATCGTCATGAGCTTGTGACGGGGGGGCACTTCTGA	CCTTCCTCCG	694
QY	1737	CACGGAGGGGGCCCCGCTGCGGGGTGAACA	CTTGCTGCAAGTGTGGGGGATGCAAGCTGC	179
Db	693	CACGAGGGGGCCCCG-CTGCGGGGTGAACA	CTTGCTGCAAGTGTGGGGGATGCAAGCTGC	635
QY	1797	TGCGATGGAGTACCTGAGAGCAAGTGTGATCCAC-C	-GGGACCTGCGTGTCCGAAT	185
Db	634	TGCGATGGAGTACCTGAGAGCAAGTGTGATCCAC	CAAGGACCTGCGTGTCCGAAT	575
QY	1856	GCCCTGTGACAGAGAAAGATGCTCTGAAGATCA	GTGACTTTGGGATGTCCCGAGAGA-A	191
Db	574	GCCCTGTGACAGAGAAAGATGCTCTGAAGATCA	GTGACTTTGGGATGTCCCGAGAGAGA	515
QY	1915	GGCGATGGGGTCTATGCAAGCTTCAAGGGGGGCTCA	ACAAGTCCCCCTGAAGTGAACCGCA	197
Db	514	GGCGATGGGGTCTATGCAAGCTTCAAGGGGGGCTCA	ACAAGTCCCCCTGAAGTGAACCGCA	455
QY	1975	CCTGAGGCCCTTAACTACGCGCGCTACTCTCTCCGAAACGAG	AGTGTGAGCTTTGGCAATC	203
Db	454	CCTGAGGCCCTTAACTACGCGCGCTACTCTCTCCGAAACGAG	AGTGTGAGCTTTGGCAATC	395
QY	2035	TTGCTTCTGGAGACCTTCAGCCTTGGGGGCTTCCCTCTAT	CCCACTTCAAGCAATCAGCAG	209
Db	394	TTGCTTCTGGAGACCTTCAGCCTTGGGGGCTTCCCTCTAT	CCCACTTCAAGCAATCAGCAG	335
QY	2095	ACAAGGGAGGTTTGTGAGAGAGGGGGGCGGTCTGCGCCTCGAC	CACTGTGTCTCATATGCC	215
Db	334	ACAAGGGAGGTTTGTGAGAGAGGGGGGCGGTCTGCGCCTCGAC	CACTGTGTCTCATATGCC	275
QY	2155	GTGTTCAAGGCTCATGAGACAGTGTCTGGGCTTATGAG	CCTGGGCAACGGCCCAAGCTTCACG	221
Db	274	GTGTTCAAGGCTCATGAGACAGTGTCTGGGCTTATGAG	CCTGGGCAACGGGCBACAGCTTCACG	215
QY	2215	ACCATCTTACCAAGAGCTGCAAGCATCTCGAAAGCGGCAT	TCGG	2256
Db	214	ACCATCTTACCAAGAGCTGCAAGCATCTCGAAAGCGGCAT	TCGG	173

RESULT 15					
BX327963					
LOCUS	BX327963	973 bp	mRNA	linear	EST 07-APR-2006
DEFINITION	BX327963 Homo sapiens PLACENTA CT 25-NORMALIZED Homo sapiens cDNA clone CS001069x21 5-PRIME, mRNA sequence.				

	ACCESSION	BX327963
	VERSION	BX327963.1
	KEYWORDS	GI:30336708
	SOURCE	EST.
	ORGANISM	Homo sapiens (human)
	AUTHORS	Homo sapiens
	TITLE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	REFERENCE	1 (base 1 to 973)
		Li, W.-B., Gruber, C., Jesssee, J. and Polayes, D.
		Full-length cDNA libraries and normalization
		Unpublished (2001)
		Contact: Genoscope
		Genoscope - Centre National de Sequencage
		2 rue Gasson Creteilux, CP 5706 - 91057 Evry cedex - FRANCE
		Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
		1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers
		end enriched, double-strand cDNA was digested with Not I and cloned
		into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
		was normalized. Library was constructed by life technologies, a
		division of Invitrogen. This sequence belongs to sequence cluster
		7663.f
		For more information about this cluster, see
		http://www.genoscope.cns.fr/cdna?c=OSBAG035ZC0_CSN03331_1ec=7663.f
	FEATURES	.
	Source	Location/Qualifiers
		1..973
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone_id="CSODI069YK21"
		/tissue_type="PLACENTA COT 25-NORMALIZED"
		/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
		/note="1st strand cDNA was primed with a NotI-oligo(dT)
		primer. Five prime end enriched, double-strand cDNA was
		digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN		
	Query Match	29.7%; Score 669.4; DB 5; Length 973;
	Best Local Similarity	98.7%; Pred. No. 1.Be-110;
	Matches 695; Conservative	0; Mismatches 7; Indels 2; Gaps 2;
Dy	619	GGCCTGTCGGCGTCACTGCAGG-ACTGCCACGAGAGATGGCTTGATCTGAAGAAGAT 677
Db	1	GGCTGTCTGGGTAACTCGCAGGNACTGCACGAGAGATGGCTTGATCTGAAGAAGAT 60
Dy	678	CTGCGAGGAATTACTTGGAGATTAGCAGCCGTGGTCAGAGATGGTGGCATTTACC 737
Db	61	CTGCGAGGAATTACTTGGAGATTAGCAGCCGTGGTCAGAGATGGTGGCATTTACC 120
Dy	738	GGAGATGGCTGCAGCTCTCCCGCATCCAGCCTGAGGGCTGAGTAACAAGGCTTCTGCG 797
Db	121	GGAGATGGCTGCAGCTCTCTCCCGCATCCAGCCTGAGGGCTGAGTAACAAGGCTTCTGCG 180
Dy	798	ACAGTATGGTCCGCACTGACGTCGCCACCCCTGTGTACGTTGATGATGATCACTGCTTGA 857
Db	181	ACAGTATGGTCCGCACTGACGTCGCCACCCCTGTGTACGTTGATGATGATCACTGCTTGA 240
Dy	858	GGAGGTGGAACCGCTGGAGCCTGGGGAGCTTCCAGCTGAAGAAGCTGACTGGAGAGAGGT 917
Db	241	GGAGGTGGAACCGCTGGAGCCTGGGGAGCTTCCAGCTGAAGAAGCTGACTGGAGAGAGGT 300
Dy	918	GCAACACACGCTGACCTCAGTGAACAAGTAGCTGGCTGGCCACCGAAGATGGTGTTCAG 977
Db	301	GCAACACACGCTGACCTCAGTGAACAAGTAGCTGGCTGGCCACCGAAGATGGTGTTCAG 360
Dy	978	GCGCAGAGATGGTTACGAGCTGCAACAGAACTCCGAAATGAAGAAGAAACCCCA 1037
Db	361	GCGCAGAGATGGTTACGAGCTGCAACAGAACTCCGAAATGAAGAAGAAACCCCA 420
Dy	1038	CCCCCGGAGCGGCTGACGCTGCTGGGCAAGAAGCAAAGTCTGTGAAGAAGCACTGCAGGG 1097
Db	421	CCCCCGGAGCGGCTGACGCTGCTGGGCAAGAAGCAAAGTCTGTGAAGAAGCACTGCAGGG 480

[illegible]

Query Match	29.7%;	Score 669.4;	DB 5;	Length 973;
Best Local Similarity	98.7%;	Pred. No. 1.8e-110;		
Matches 695;	Conservative	0;	Mismatches 7;	Indels 2;
Gaps	2;			
Qy 619	GGCGTCGTGGGTCGTCACGTGACAG--ACCTGGCACGAGGATGGCTTCATCTCGAAGAGAT	677		
Db 1	GGCGTCGTGGGTCGTCACGTGACAG--ACCTGGCACGAGGATGGCTTCATCTCGAAGAGAT	60		
Qy 678	CTGCGAGAAATACCTGGAGATTAGCAGCGCTGGTCAGAGATAGGTGGCCATTCAACCG	737		
Db 61	CTGCGAGAAATACCTGGAGATTAGCAGCGCTGGTCAGAGATAGGTGGCCATTCAACCG	120		
Qy 738	GGAGATGGCTGCAAGCTCTGCGCCGCATCCAGCCTGAGGCTGAGTACCAAGGCTTCTGCG	797		
Db 121	GGAGATGGCTGCAAGCTCTGCGCCGCATCCAGCCTGAGGCTGAGTACCAAGGCTTCTGCG	180		
Qy 798	ACAGTATGGGTCGCGCACTGACGCTCCACCCCTGGTCAAGCTTCGATGATGATCACTGCTTGA	857		
Db 181	ACAGTATGGGTCGCGCACTGACGCTCCACCCCTGGTCAAGCTTCGATGATGATCACTGCTTGA	240		
Qy 858	GGAGGAGTGAACCGCTGGAGCCTGGGGAGCTTCAGACTGAACGAGCTGACTGTGAGAGCGT	917		
Db 241	GGAGGAGTGAACCGCTGGAGCCTGGGGAGCTTCAGACTGAACGAGCTGACTGTGAGAGCGT	300		
Qy 918	GCAGCAACCGCTACCTCAAGTCAAGATGAGCTGGCTGGCCACCGAGATGGTGTTCAG	977		
Db 301	GCAGCAACCGCTACCTCAAGTCAAGATGAGCTGGCTGGCCACCGAGATGGTGTTCAG	360		
Qy 978	GCGGCAGAGATGGTTACGCGAGCTGCAACAGAGACTTCGGATATGAAAGAGAACACCA	1037		
Db 361	GCGGCAGAGATGGTTACGCGAGCTGCAACAGAGACTTCGGATATGAAAGAGAACACCA	420		
Qy 1038	CCCCCGGAGCGGCTGCACTGCTGGGCGCAAGAGCAAGTGTGCAAGAGCACTGCAAGG	1097		
Db 421	CCCCCGGAGCGGCTGCACTGCTGGGCGCAAGAGCAAGTGTGCAAGAGCACTGCAAGG	480		

Query Match	29.7%;	Score 669.4;	DB 5;	Length 973;
Best Local Similarity	98.7%;	Pred. No. 1.8e-110;		
Matches 695;	Conservative	0;	Mismatches 7;	Indels 2;
Gaps	2;			
Qy 619	GGCGTCGTGGGTCGTCACGTGACAG--ACCTGGCACGAGGATGGCTTCATCTCGAAGAGAT	677		
Db 1	GGCGTCGTGGGTCGTCACGTGACAG--ACCTGGCACGAGGATGGCTTCATCTCGAAGAGAT	60		
Qy 678	CTGCGAGAAATACCTGGAGATTAGCAGCGCTGGTCAGAGATAGGTGGCCATTCAACCG	737		
Db 61	CTGCGAGAAATACCTGGAGATTAGCAGCGCTGGTCAGAGATAGGTGGCCATTCAACCG	120		
Qy 738	GGAGATGGCTGCAAGCTCTGCGCCGCATCCAGCCTGAGGCTGAGTACCAAGGCTTCTGCG	797		
Db 121	GGAGATGGCTGCAAGCTCTGCGCCGCATCCAGCCTGAGGCTGAGTACCAAGGCTTCTGCG	180		
Qy 798	ACAGTATGGGTCGCGCACTGACGCTCCACCCCTGGTCAAGCTTCGATGAGTCACTGCTTGA	857		
Db 181	ACAGTATGGGTCGCGCACTGACGCTCCACCCCTGGTCAAGCTTCGATGAGTCACTGCTTGA	240		
Qy 858	GGAAGGTGAACCGCTGGAGCCTGGGGAGCTTCAGACTGAACGAGCTGACTGTGAGAGCGT	917		
Db 241	GGAAGGTGAACCGCTGGAGCCTGGGGAGCTTCAGACTGAACGAGCTGACTGTGAGAGCGT	300		
Qy 918	GCAGCAACCGCTACCTCAAGTCAAGATGAGCTGGCTGGCCACCGAGATGGGTTCAG	977		
Db 301	GCAGCAACCGCTACCTCAAGTCAAGATGAGCTGGCTGGCCACCGAGATGGGTTCAG	360		
Qy 978	GCGGCAGAGATGGTTACGCGAGCTGCAACAGAGACTTCGGATGTAAGAGAGAACCCCA	1037		
Db 361	GCGGCAGAGATGGTTACGCGAGCTGCAACAGAGACTTCGGATGTAAGAGAGAACCCCA	420		
Qy 1038	CCCCCGGAGCGGCTGCACTGCTGGGCGCAAGAGCAAGTGTGCAAGAGCACTGCAAGG	1097		
Db 421	CCCCCGGAGCGGCTGCACTGCTGGGCGCAAGAGCAAGTGTGCAAGAGCACTGCAAGG	480		

Qy 1098 GCTGCAAGTACGCGCTGTGTCAGCCAGGCAAGCTGCAGAGCCACAGAGATTGCTGCAGAC 1157
Db 481 GCTGCAAGTACGCGCTGTGTCAGCCAGGCAAGCTGCAGAGCCACAGAGATTGCTGCAGAC 540
Qy 1158 CAAGCTGAGACACTGAGGCCCCCGGCGAGAGCCCGCTGTGCTGCTCTCTGAGATGACCG 1217
Db 541 CAAGCTGAGACACTGAGGCCCCCGGCGAGAGCCCGCTGTGCTGCTCTCTGAGATGACCG 600
Qy 1218 CCACTCCACGTCGTCTCGGAGCAGAGAGAGAGAGG--GGGAGGACACCCACGCTGAGAG 1276
Db 601 CCACTCCACGTCGTCTCGGAGCAGAGAGAGAGAGACCCACGCTGAGAG 660
Qy 1277 TCCTTAAGAGCAATCTCAGAAATCTTCCGCCCAAGTTCTCG 1320
Db 661 TCCTTAAGAGCAATCTCAGAAATCTTCCGCCCAAGTTCTCG 704

Search completed: March 19, 2005, 14:10:46
Job time : 6521.86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 01:20:31 ; Search time 1076.29 Seconds
(without alignment)
12408.301 Million cell updates/sec

Title: US-10-660-763-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 atggcctctctctctgcgact.....gcattccgaagcgcgcacg 2256

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16dec04:*
1: geneseqn19808:*
2: geneseqn19808:*
3: geneseqn20008:*
4: geneseqn2001as:*
5: geneseqn2001as:*
6: geneseqn2002as:*
7: geneseqn2002as:*
8: geneseqn2003as:*
9: geneseqn2003as:*
10: geneseqn2003as:*
11: geneseqn2003as:*
12: geneseqn2004as:*
13: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2256	100.0	2674	6	ABA94500 Human pro
2	2256	100.0	2674	8	ABV75413 Human kin
3	2256	100.0	2674	12	ADM28580 Human pro
4	2206.2	97.8	2599	12	ADK71888 Human kin
5	2023.4	89.7	2689	6	AA895001 Human DNA
6	1671.6	74.1	2623	13	ADR39816 Human kin
7	1540.4	68.3	2680	12	ADH43089 CRAM prot
8	635	28.1	856	1	AAAT70060 fcs/fps p
9	375	16.6	2818	11	ADL71052 Gene enco
10	348.8	15.5	527	13	ADQ54298 Novel can
11	328.8	14.6	2584	11	ADL71054 Gene enco
12	326.8	14.5	449	9	ACH14291 Human adu
13	296.2	13.1	361	6	AA817052 Human pro
14	296.2	13.1	361	10	AA61604 Human pro
15	296.2	13.1	361	10	ABX14976 Human pro
16	271.6	12.0	1779	8	ACC57665 Drosophi
17	271.6	12.0	3198	4	AB118403 Drosophi
18	271.6	12.0	3369	4	AB114457 Drosophi
19	227.2	10.1	15297	6	ABA94501 Human pro
20	227.2	10.1	15297	8	ABV75414 Human kin

21	227.2	10.1	15297	12	ADM28582 Human pro
22	225.6	10.0	12263	6	ABK84514 Human cdy
23	201.2	8.9	3286	12	ADK71885 Human kin
24	201.2	8.9	3875	6	ABZ35414 Human gen
25	199.6	8.8	5616	10	AD608824 Novel DNA
26	199.2	8.8	2778	12	ADP28218 Human sec
27	198	8.8	2955	6	ABV78117 Human eph
28	198	8.8	2955	6	ABZ35693 Human eph
29	198	8.8	2955	6	ABX09936 Human eph
30	198	8.8	2955	6	AB191658 Human pol
31	198	8.8	3370	6	ABL67952 Ovary can
32	198	8.8	3370	8	ABZ34861 Coding se
33	198	8.8	3370	10	ADD18447 Human pro
34	198	8.8	3370	10	ADP38440 Human pro
35	198	8.8	3370	12	ADJ75094 Marker ge
36	198	8.8	3370	13	ADQ80244 Eph1 cDNA
37	185.4	8.2	3042	6	ABV78118 Human eph
38	185.4	8.2	3042	6	ABZ35694 Human eph
39	185.4	8.2	3042	6	ABX09937 Human eph
40	185.4	8.2	3042	6	ABZ91659 Human pol
41	185.4	8.2	3346	13	ACN42099 Human dia
42	185.4	8.2	3319	13	ACN37278 Tumour-as
43	185.4	8.2	3921	6	ABV94231 Breast ca
44	185.4	8.2	3921	8	ABZ34865 Coding se
45	185.4	8.2	3921	10	ADB37439 Human can

ALIGNMENTS

RESULT 1	ABV75413
ID	ABA94500 standard; cDNA; 2674 BP.
XX	ABA94500;
AC	09-APR-2002 (first entry)
XX	
DT	Human proto-oncogene tyrosine kinase encoding cDNA.
XX	
XX	Proto-oncogene tyrosine kinase; potK; tumour; cytolethic; anti-leukemic;
KW	gene therapy; protein therapy; vaccine; enzymatic-inhibition; human;
KW	anti-kinase; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	5'UTR
FT	Location/Qualifiers
FT	1..71
FT	/*tag= a
FT	72..2330
FT	/*tag= b
FT	/product= "proto-oncogene tyrosine kinase"
FT	2331..2674
FT	/*tag= c
XX	
XX	3'UTR
XX	
XX	US6340584-B1.
XX	
XX	22-JAN-2002.
XX	
XX	27-MAR-2001; 2001US-00817180.
XX	
XX	27-MAR-2001; 2001US-00817180.
XX	
XX	(PEKE) PE CORP NY.
XX	
XX	Gan W, Ye J, Di Francesco V, Beasley EM;
XX	WPI, 2002-138497/18.
XX	P-PDSB; ABB07354.
XX	
XX	Nucleic acids encoding a proto-oncogene tyrosine kinase, useful for the
XX	prevention, diagnosis and treatment of e.g. leukemia and lung tumors.

PS Claim 1; Fig 1A-B; 49p; English.

XX The invention provides isolated nucleic acid sequences encoding a proto-oncogene tyrosine kinase (pork). The pork polynucleotides and protein may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate pork expression, such as lung and kidney tumors, leukemia and stomach adenocarcinoma. pork may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of pork by expressing inactive proteins or to supplement the patient's own production of pork. The encoded pork may be used as an antigen in the production of pork antibodies against pork and in assays to identify modulators of pork expression and activity. The anti-pork antibodies and antagonists may be used to down regulate expression and activity and as diagnostic agents for detecting the presence of pork in samples. The present sequence represents a cDNA encoding the human pork

XX Sequence 2674 BP; 662 A; 767 C; 810 G; 435 T; 0 U; 0 Other;

SQ Query Match 100.0%; Score 2256; DB 6; Length 2674;

Best Local Similarity 100.0%; Pred. No. 0; Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGCTTCTTCTGAGCTGTGACGCCCCCAGGACGAGGCTCTTGACGCAATGACG 60
DB 72 ATGGGCTTCTTCTGAGCTGTGACGCCCCCAGGAGCCAGGGGTCTTGACGCAATGACG 131
QY 61 GAGGCGGACCTTCTGCTACTGAGGGGATGAAAGTGAATGATGCCAGGGGGTCAAGAT 120
DB 132 GAGGCGGACCTTCTGCTACTGAGGGGATGAAAGTGAATGATGCCAGGGGGTCAAGAT 191
QY 121 GACAGGAGATATGACAGACTGCTTACCAATGTCCCTGACAGGACAGTGGGGCCAGAGC 180
DB 192 GACAGGAGATATGACAGACTGCTTACCAATGTCCCTGACAGGACAGTGGGGCCAGAGC 251
QY 181 CGGGCCATCAGGCTTGACAGGCCCCATCAGTCACTGTGGCTGAGATCAACGCAAACT 240
DB 252 CGGGCCATCAGGCTTGACAGGCCCCATCAGTCACTGTGGCTGAGATCAACGCAAACT 311
QY 241 GAGGGCTTACGCGCTTGTGCGGGGACGACGCGAGAGATCTGAATCAAGAGGCCCTGAC 300
DB 312 GAGGGCTTACGCGCTTGTGCGGGGACGACGCGAGAGATCTGAATCAAGAGGCCCTGAC 371
QY 301 AAGCTGAGCTGCTGCTACCGGGGAAACGAGCAGCTTCCGAAAGACTACAGCAGCAGTGG 360
DB 372 AAGCTGAGCTGCTGCTACCGGGGAAACGAGCAGCTTCCGAAAGACTACAGCAGCAGTGG 431
QY 361 CAGCAGCTCAGAGAGGCTCACCAAGACCCAGCCAGGACATTTGAAAGCTGAAGAC 420
DB 432 CAGCAGCTCAGAGAGGCTCACCAAGACCCAGCCAGGACATTTGAAAGCTGAAGAC 491
QY 421 CAGTACCGAGCTTGACAGGGAAGTGTCCAAAGCCAAAGGCAAGTACAGAGAGGCGAC 480
DB 492 CAGTACCGAGCTTGACAGGGAAGTGTCCAAAGCCAAAGGCAAGTACAGAGAGGCGAC 551
QY 481 AAAGACAAGAGCCGTGACAAAGCCAAAGCAATATATGTGGCAGCTGTGAAAGCTTT 540
DB 552 AAAGACAAGAGCCGTGACAAAGCCAAAGCAATATATGTGGCAGCTGTGAAAGCTTT 611
QY 541 GCTCACCACAAACGCTATGTGCTGGGCGTGGGCTTGCGAGCTACCAACCAACGACAC 600
DB 612 GCTCACCACAAACGCTATGTGCTGGGCGTGGGCTTGCGAGCTACCAACCAACGACAC 671
QY 601 CACCAACTCTGTGCTCCCGGCTGCTGCGGTCACTGCAAGACCTTGACAGAGAGATGGCT 660
DB 672 CACCAACTCTGTGCTCCCGGCTGCTGCGGTCACTGCAAGACCTTGACAGAGAGATGGCT 731
QY 661 TGCATCTGAAAGAGATCTGAGAAATACCTGGAATTAAGAGCTGTGAGATGAG 720
DB 732 TGCATCTGAAAGAGATCTGAGAAATACCTGGAATTAAGAGCTGTGAGATGAG 791
QY 721 GTGTGAGCAATTCACGGGAGATGAGTGCAGCTGTGCGGCAATCCAGGCTGAGGCTGAG 780

DB 792 GTGTGAGCAATTCACGGGAGATGAGTGCAGCTGTGCGGCAATCCAGGCTGAGGCTGAG 851
QY 781 TACCAAGGCTTCTGCGACAGTATGAGGTCCGACCTGACGTCCACCCCTGTGACGTTTC 840
DB 852 TACCAAGGCTTCTGCGACAGTATGAGGTCCGACCTGACGTCCACCCCTGTGACGTTTC 911
QY 841 GATGAGTCACTGCTTGAAGAGGAGTGAACGCTGAGGCTGGGAGGCTCCAGCTGAACGAG 900
DB 912 GATGAGTCACTGCTTGAAGAGGAGTGAACGCTGAGGCTGGGAGGCTCCAGCTGAACGAG 971
QY 901 CTGACTGTGAGAGAGGTGACGACACGCTGACCTTACATGACAGATGAGTGTGCTGGCC 960
DB 972 CTGACTGTGAGAGAGGTGACGACACGCTGACCTTACATGACAGATGAGTGTGCTGGCC 1031
QY 961 ACCGAGATGAGTGTGAGGCGGAGAGAGATGTTTACAGCTGCAACAGAGGCTCCGGAAT 1020
DB 1032 ACCGAGATGAGTGTGAGGCGGAGAGAGATGTTTACAGCTGCAACAGAGGCTCCGGAAT 1091
QY 1021 GAAAGAGAAACACCAACCCCGGAGCGGATGACGCTGTGGGCAAGAGCAATGCTTG 1080
DB 1092 GAAAGAGAAACACCAACCCCGGAGCGGATGACGCTGTGGGCAAGAGCAATGCTTG 1151
QY 1081 CAAAGACATGACAGGGGCTGAGAGTATGCTGTGACGACGACCCCAAGCTGACAGGCCAG 1140
DB 1152 CAAAGACATGACAGGGGCTGAGAGTATGCTGTGACGACGACCCCAAGCTGACAGGCCAG 1211
QY 1141 CAGAGTGTGCTGACAGCAACCAAGCTGAGACCTGTGGGCCCCGCGCTGTGCTG 1200
DB 1212 CAGAGTGTGCTGACAGCAACCAAGCTGAGACCTGTGGGCCCCGCGCTGTGCTG 1271
QY 1201 CTCTGACAGATGACCGCCACTTCAAGTGTCTCTGGAACAGAGAGAGAGAGGAGAG 1260
DB 1272 CTCTGACAGATGACCGCCACTTCAAGTGTCTCTGGAACAGAGAGAGAGAGGAGAG 1331
QY 1261 AACCCACGCTGAGATCTTAAAGCCCAATCTCAGAAATCTTCCGCCCAAGTTCTG 1320
DB 1332 AACCCACGCTGAGATCTTAAAGCCCAATCTCAGAAATCTTCCGCCCAAGTTCTG 1391
QY 1321 AACCTGTACCACTGGAAGAGGAGGAGGCTTCTTACAGATTCCTTGTGCTACACCACTA 1380
DB 1392 AACCTGTACCACTGGAAGAGGAGGAGGCTTCTTACAGATTCCTTGTGCTACACCACTA 1451
QY 1381 CTGAGACCCAGCAGACCCCTCTCAAAAGAGTGTGTCTGTGACAGGGCTGTGCC 1440
DB 1452 CTGAGACCCAGCAGACCCCTCTCAAAAGAGTGTGTGTCTGTGACAGGGCTGTGCC 1511
QY 1441 AAGACAAGTGGGTCTGAACCAATGAGACCTGTGTGAGGAGAGATTTGAACGGGG 1500
DB 1512 AAGACAAGTGGGTCTGAACCAATGAGACCTGTGTGAGGAGAGATTTGAACGGGG 1571
QY 1501 AACTTGGGGAAGTGTTCAGCGGAGCGCTGCGAGCGCAACAACCTGTGGCGGTGAAG 1560
DB 1572 AACTTGGGGAAGTGTTCAGCGGAGCGCTGCGAGCGCAACAACCTGTGGCGGTGAAG 1631
QY 1561 TCTTGTGAGAGACGCTCCACCTGACCTCAAGAGCCAAAGTTCACAGAAAGCAGATTC 1620
DB 1632 TCTTGTGAGAGACGCTCCACCTGACCTCAAGAGCCAAAGTTCACAGAAAGCAGATTC 1691
QY 1621 CTGAAGCAGTACAGCACCACCAATCGTGCCTTCAATTTGTGTCTGCAACCCAGAGCAG 1680
DB 1692 CTGAAGCAGTACAGCACCACCAATCGTGCCTTCAATTTGTGTCTGCAACCCAGAGCAG 1751
QY 1681 CCATCTACATGATGATGAGCTTGTGACGGGGGGGCACTTCTGACCTTCTCCGCAAG 1740
DB 1752 CCATCTACATGATGATGAGCTTGTGACGGGGGGGCACTTCTGACCTTCTCCGCAAG 1811
QY 1741 GAGGGGCCCCGCTGCGGATGAAGACTGTGCTGACAGATGTGGGAGTCAAGCTGTGCG 1800
DB 1812 GAGGGGCCCCGCTGCGGATGAAGACTGTGCTGACAGATGTGGGAGTCAAGCTGTGCG 1871
QY 1801 ATGAGATCTTGAAGAGCAAGTGTGATCAACCGGAGACCTGTGCTGTGGAATCTGCTG 1860
DB 1872 ATGAGATCTTGAAGAGCAAGTGTGATCAACCGGAGACCTGTGCTGTGGAATCTGCTG 1931


```

Db      612 GCTCACCAACCGCTATGCTGCTGGGCTGCGGGCTGCGACGACGACCAACGACACAC 671
Qy      601 GACCAAGCTCTGCTGCCCCGGCCCTGCTGGGTGCTGCAAGAGAGAGAGAGAGAGT 660
Db      672 CACCAAGCTCTGCTGCCCCGGCCCTGCTGGGTGCTGCAAGAGAGAGAGAGAGAGT 731
Qy      661 TGCATCTGAAGAGAGATCTGCAAGATACCTGAGAGATTAGCAGCTGTGTGAGATGAG 720
Db      732 TGCATCTGAAGAGAGATCTGCAAGATACCTGAGAGATTAGCAGCTGTGTGAGATGAG 791
Qy      721 GTGTGGCCATTCACCGGAGAGATGCTGCAAGCTGTGCTCCCGACATCCAGCTTGAAG 780
Db      792 GTGTGGCCATTCACCGGAGAGATGCTGCAAGCTGTGCTCCCGACATCCAGCTTGAAG 851
Qy      781 TACCAAGGCTTCCCTGGAGAGATGAGTGGGTCCGACCTGAGAGTCCACCCCTGTGCAAGTTC 840
Db      852 TACCAAGGCTTCCCTGGAGAGATGAGTGGGTCCGACCTGAGAGTCCACCCCTGTGCAAGTTC 911
Qy      841 GATGAGTCACTGCTTGAAGAGAGGATGAAACGCTGAGAGCTTGGGAGCTCCAGCTGAACGAG 900
Db      912 GATGAGTCACTGCTTGAAGAGAGGATGAAACGCTGAGAGCTTGGGAGCTCCAGCTGAACGAG 971
Qy      901 CTGACTGTGAGAGAGCTGAGACACACGCTGACCTGAGACAGATGAGCTGTGGCTGACC 960
Db      972 CTGACTGTGAGAGAGCTGAGACACACGCTGACCTGAGACAGATGAGCTGTGGCTGACC 1031
Qy      961 ACCGAGATGCTGTTCAAGGCGGAGAGATGTTACGACCTGACAGAGAGAGCTCCGGAAT 1020
Db      1032 ACCGAGATGCTGTTCAAGGCGGAGAGATGTTACGACCTGACAGAGAGAGCTCCGGAAT 1091
Qy      1021 GAAAGAGAGAAACACCCACCCCGGAGCGGAGTCACTGCTGGGCAAGAGGCAAGTGTCTG 1080
Db      1092 GAAAGAGAGAAACACCCACCCCGGAGCGGAGTCACTGCTGGGCAAGAGGCAAGTGTCTG 1151
Qy      1081 CAAGAAGCACTGAGGGGCTGCAAGTACGCTGTGACAGCCAGGCTGACGAGCCAG 1140
Db      1152 CAAGAAGCACTGAGGGGCTGCAAGTACGCTGTGACAGCCAGGCTGACGAGCCAG 1211
Qy      1141 CAGGAGTTCTGAGAGCAAGCTGAGAGCACTGGGGCCCGGCGAGGCCCCGCTGTGCTG 1200
Db      1212 CAGGAGTTCTGAGAGCAAGCTGAGAGCACTGGGGCCCGGCGAGGCCCCGCTGTGCTG 1271
Qy      1201 CTCCTGCAAGATGACCGGCACTCCAGCTGCTCTCGAGACAGAGAGAGAGGAGGAG 1260
Db      1272 CTCCTGCAAGATGACCGGCACTCCAGCTGCTCTCGAGAGAGAGAGAGGAGGAG 1331
Qy      1261 ACAACCAAGCTGAGATCTTAAAGACCAATCTGAGAAATCTTCCGCCCAAGTTCTCG 1320
Db      1332 ACAACCAAGCTGAGATCTTAAAGACCAATCTGAGAAATCTTCCGCCCAAGTTCTCG 1391
Qy      1321 AACCTGTACCGATGGAAGGGGAGGCTTCTAGCAATCTTGTGCTCATCCAGCACTTA 1380
Db      1392 AACCTGTACCGATGGAAGGGGAGGCTTCTAGCAATCTTGTGCTCATCCAGCACTTA 1451
Qy      1381 CTGAGACACCAAGAGCCCTCCTACCAAGAAAGTGTGTCTCTGCAAGAGGCTGTGCC 1440
Db      1452 CTGAGACACCAAGAGAGCCCTCCTACCAAGAAAGTGTGTCTCTGCAAGAGGCTGTGCC 1511
Qy      1441 AAGGACAAAGTGGTGTCTGAAACATGAGAGACCTGTGTGTGGTGAAGCAATTTGAGCGGG 1500
Db      1512 AAGGACAAAGTGGTGTCTGAAACATGAGAGACCTGTGTGTGGTGAAGCAATTTGAGCGGG 1571
Qy      1501 AACTTTGGGCAAGTGTTCACGGGAGCGCTGCGAGCGGAGCAACACCTGTGGGGGGTGAAG 1560
Db      1572 AACTTTGGGCAAGTGTTCACGGGAGCGCTGCGAGCGGAGCAACACCTGTGGGGGGTGAAG 1631
Qy      1561 TCTTGTGAGAGAGGCTCCCACTGACCTCAAGGCAAGTTTCTAAGAGAAAGGAGATC 1620
Db      1632 TCTTGTGAGAGAGGCTCCCACTGACCTCAAGGCAAGTTTCTAAGAGAAAGGAGATC 1691
Qy      1621 CTGAGACAGTACAGCAACCCCAACATCTGTGCTCTCATTTGGTGTGCAACCCAGAACAG 1680
Db      1692 CTGAGACAGTACAGCAACCCCAACATCTGTGCTCTCATTTGGTGTGCAACCCAGAACAG 1751

```

```

Qy      1681 CCCATCTACATGCTATGAGAGTTGTGACAGGGGGGAGATTCTCTGACCTTCTCCGACG 1740
Db      1752 CCCATCTACATGCTATGAGAGTTGTGACAGGGGGGAGATTCTCTGACCTTCTCCGACG 1811
Qy      1741 GAGGGGGCCCGCTCGGGTGAAGA CTGCTGCAAGATGTTGGGGATGACAGCTGTGTC 1800
Db      1812 GAGGGGGCCCGCTCGGGTGAAGA CTGCTGCAAGATGTTGGGGATGACAGCTGTGTC 1871
Qy      1801 ATGAGATACCTTGAAGAGCAAGTGTGATCCACCCGGAGACCTGCTGTGAAATGCCCTG 1860
Db      1872 ATGAGATACCTTGAAGAGCAAGTGTGATCCACCCGGAGACCTGCTGTGAAATGCCCTG 1931
Qy      1861 GTGACAGAGAAATGTCCTTGAAGATCAAGTGAATTTGGAGTGTCCGAGAGAGAGCCGAT 1920
Db      1932 GTGACAGAGAAATGTCCTTGAAGATCAAGTGAATTTGGAGTGTCCGAGAGAGAGCCGAT 1991
Qy      1921 GGGGTCTATGACAGCTCAAGGGGGCTTCAGACAAAGTCCCGTGAAGTGAACCGACCTGAG 1980
Db      1992 GGGGTCTATGACAGCTCAAGGGGGCTTCAGACAAAGTCCCGTGAAGTGAACCGACCTGAG 2051
Qy      1981 GCCCTTAATACAGGCGCTACTCTCCGAAACCGAGAGTGTGAGCTTTGGCAATCTTGTCTC 2040
Db      2052 GCCCTTAATACAGGCGCTACTCTCCGAAACCGAGAGTGTGAGCTTTGGCAATCTTGTCTC 2111
Qy      2041 TGGAGAACCTTCAAGCTGAGGGGCTCCCGCTAATCCCAACCTGACAGCAATCAGACACCG 2100
Db      2112 TGGAGAACCTTCAAGCTGAGGGGCTCCCGCTAATCCCAACCTGACAGCAATCAGACACCG 2171
Qy      2101 GAGTTTGTGAGAGAGGGGGCCGTCTGCTCCCAAGAGCTGTCTGATGCCGTGTC 2160
Db      2172 GAGTTTGTGAGAGAGGGGGCCGTCTGCTCCCAAGAGCTGTCTGATGCCGTGTC 2231
Qy      2161 AAGCTCATGAGAGCAATGCTGTGGGCTATGAGCTGTGGGAGCGGAGCCCACTTCAGACACATC 2220
Db      2232 AAGCTCATGAGAGCAATGCTGTGGGCTATGAGCTGTGGGAGCGGAGCCCACTTCAGACACATC 2291
Qy      2221 TACCAAGAGCTGACAGAGATCCGAAAGCGGCAATCG 2256
Db      2292 TACCAAGAGCTGACAGAGATCCGAAAGCGGCAATCG 2327

RESULT 3
ADM28580
ID      ADM28580   Brandard; cDNA; 2674 BP.
XX      AC          ADM28580;
XX      DT          17-JUN-2004 (first entry)
XX      DE          Human proto-oncogene tyrosine kinase cDNA.
XX      KW          ss; gene; gene therapy; human; proto-oncogene; tyrosine kinase; cancer.
XX      OS          Homo sapiens.
XX      FH          Key
FT          5'UTR          1..71
FT          CDS            /*tag= a
FT          FT            72..2330
FT          FT            /*tag= b
FT          FT            /product= "Proto-oncogene tyrosine kinase"
FT          FT            2331..2674
FT          FT            /*tag= c
XX      PN          US2004063130-A1.
XX      PD          01-APR-2004.
XX      PF          12-SEP-2003; 2003US-00660763.
XX      PR          27-MAR-2001; 2001US-00817180.
XX      PR          06-DEC-2001; 2001US-00003295.

```


DB 515 CAGTACCGAGCTCTGSCACGGGACACTGCCCCAACCGCAAGTACAGAGAGGCGAC 574
QY 481 AAGAGACAGAGCCGTGACAAAGCCAAAGAACAGATATGTGGAGCGCTGTGAAAGCTCTTT 540
DB 575 AAGAGACAGAGCCGTGACAAAGCCAAAGAACAGATATGTGGAGCGCTGTGAAAGCTCTTT 634
QY 541 GGTCAACCAACCGCTATGTGTGGGGCGTGGGGCTGGCCAGTTAACACCAACAGCAAC 600
DB 635 GCTCACCAACCGCTATGTGTGGGGCGTGGGGCTGGCCAGTTAACACCAACAGCAAC 694
QY 601 CACAGAGCTCTGCTGCGCGGCTGCTGCGGCTGCTGAGGATCTGACAGGAGATGAGCT 660
DB 695 CACAGAGCTCTGCTGCGCGGCTGCTGCGGCTGCTGAGGATCTGACAGGAGATGAGCT 754
QY 661 TGCATCTGAGAGAGATCTCTGAGAGAAATCTGAGAGATTAGCAGCTGTGTGAGATGAG 720
DB 755 TGCATCTGAGAGAGATCTCTGAGAGAAATCTGAGAGATTAGCAGCTGTGTGAGATGAG 814
QY 721 GTGTGGCCATTACACCGGAGAGATGCTGAGAGCTGCTGCGGCTGCTGAGCTGAGCTGAG 780
DB 815 GTGTGGCCATTACACCGGAGAGATGCTGAGAGCTGCTGCGGCTGCTGAGCTGAGCTGAG 874
QY 781 TACCAAGGCTCTCTGCGAGAGATATGAGGCTCGCACCTGACCTGCTGTGTGACGCTTC 840
DB 875 TACCAAGGCTCTCTGCGAGAGATATGAGGCTCGCACCTGACCTGCTGTGTGACGCTTC 934
QY 841 GATGAGTCACTGCTTGAAGAGGATGAAACCGCTGAGAGCTGTGGAGGCTCCAGCTGAAAG 900
DB 935 GATGAGTCACTGCTTGAAGAGGATGAAACCGCTGAGAGCTGTGGAGGCTCCAGCTGAAAG 994
QY 901 CTGAGCTGTGAGAGGCTGACAGCAGCTGCTGAGTGAAGTGAAGTGTGTGTGCTGCTG 960
DB 995 CTGAGCTGTGAGAGGCTGACAGCAGCTGCTGAGTGAAGTGAAGTGTGTGTGCTGCTG 1054
QY 961 ACCGAGATGATGTTGACGCGGAGAGAGATATGTTAACGAGCTGACAGAGAGCTCCGGAAT 1020
DB 1055 ACCGAGATGATGTTGACGCGGAGAGAGATATGTTAACGAGCTGACAGAGAGCTCCGGAAT 1114
QY 1021 GAAGAGAGAGAACCCACCCCGGAGAGCGGATGACGCTGTGGGCAAGAGAGAGTGTCTG 1080
DB 1115 GAAGAGAGAGAACCCACCCCGGAGAGCGGATGACGCTGTGGGCAAGAGAGAGTGTCTG 1174
QY 1081 CAAAGAGAGCTGAGAGGAGCTGAGAGTATGAGGCTGTGAGAGCTGAGAGAGCTGAGAG 1140
DB 1175 CAAAGAGAGCTGAGAGGAGCTGAGAGTATGAGGCTGTGAGAGCTGAGAGAGCTGAGAG 1234
QY 1141 CAGAGAGTCTGAGAGCAAGCTGAGAGCACTGGGAGCCCGGAGAGCCCGCTGTGTCTG 1200
DB 1235 CAGAGAGTCTGAGAGCAAGCTGAGAGCACTGGGAGCCCGGAGAGCCCGCTGTGTCTG 1294
QY 1201 CTCTGAGAGATGACCGCACTTCAAGTGTCTCTC----- 1235
DB 1295 CTCTGAGAGATGACCGCACTTCAAGTGTCTCTCTGCGGCAACCCCGTGAAGTGTCTG 1354
QY 1236 -----GAG 1285
DB 1355 CTTGGCTTCAAG 1414
QY 1286 GGCACATCTCAGAGAACTTCCCGCCCAAGTCTCGAAGCTGTACCGAGCTGAGAGAGAGAG 1345
DB 1415 GGCACATCTCAGAGAACTTCCCGCCCAAGTCTCGAAGCTGTACCGAGCTGAGAGAGAGAG 1474
QY 1346 GCTTCTCTAGCAATCTTCTGTCTCATGACCACTTATGAGCAACCAAGAGAGAGAGAGAG 1405
DB 1475 GCTTCTCTAGCAATCTTCTGTCTCATGACCACTTATGAGCAACCAAGAGAGAGAGAGAG 1534
QY 1406 AAGAGAGTGTGTGTGTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1465
DB 1535 AAGAGAGTGTGTGTGTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1594
QY 1466 AAGAGAGTGTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1525

DB 1595 AAGAGCTGTGTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1654
QY 1526 GCTTCTCTAGCAATCTTCTGTCTCATGACCACTTATGAGCAACCAAGAGAGAGAGAGAG 1585
DB 1655 GCTTCTCTAGCAATCTTCTGTCTCATGACCACTTATGAGCAACCAAGAGAGAGAGAGAG 1714
QY 1586 AAGAGAGTGTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1645
DB 1715 AAGAGAGTGTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1774
QY 1646 TGTGCTCTCATGATGTGTCTGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1705
DB 1775 TGTGCTCTCATGATGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1834
QY 1706 TGTGAG 1765
DB 1835 TGTGAG 1894
QY 1766 TGTGCTCTCATGATGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1825
DB 1895 TGTGCTCTCATGATGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1954
QY 1826 GATGAG 1885
DB 1955 GATGAG 2014
QY 1886 TGAATGATCTTGTGATGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1945
DB 2015 TGAATGATCTTGTGATGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2074
QY 1946 TGAAG 2005
DB 2075 TGAAG 2134
QY 2006 CCGAAG 2065
DB 2135 CCGAAG 2194
QY 2066 CCGAAG 2125
DB 2195 CCGAAG 2254
QY 2126 TGTGCTCTCATGATGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2185
DB 2255 TGTGCTCTCATGATGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2314
QY 2186 ATGAGAGTGTGAG 2245
DB 2315 ATGAGAGTGTGAG 2374
QY 2246 AAG 2305
DB 2375 AAG 2434
RESULT 5
AAS95001
ID AAS95001 standard; DNA; 2889 BP.
XX
AC AAS95001;
XX
DT 14-FEB-2002 (first entry)
XX
XX Human DNA sequence #256 expressed during foam cell differentiation.
DE Human; foam cell differentiation; atherosclerosis; cerebral stroke;
KW cardiovascular disorder; coronary artery disease; gene therapy; ds.
OS Homo sapiens.
XX
PN WO200177389-A2.
XX
PD 18-OCT-2001.

XX 04-APR-2001; 2001WO-US011128.
XX
XX
XX 05-APR-2000; 2000US-0195106P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Shiffman D, Somogyi R, Lawn R, Selthamer JJ, Porter GJ, Mikita T,
XX Tai J,
XX WPI, 2002-010925/01.
XX
XX Composition useful for diagnosis of conditions, disorders or diseases
XX associated with atherosclerosis, comprises several polynucleotides that
XX are differentially expressed in foam cell development.
XX
XX Claim 1; Page 298-299; 315pp; English.
XX
XX The present invention relates to the isolation of human polynucleotide
XX sequences that are differentially expressed during foam cell
XX differentiation. The polynucleotide sequences of the invention or a
XX composition comprising these polynucleotides are useful as a high
XX throughput method for detecting altered expression of one or more
XX polynucleotides in a sample. The polynucleotides can be used in the
XX diagnosis of disorders associated with foam cell development such as
XX atherosclerosis, cerebral stroke, and cardiovascular disorders such as
XX coronary artery disease. The polynucleotide sequences can also be used as
XX PCR primers and probes. The polynucleotides of the invention are also
XX useful in gene therapy. AAS94746-AAS95021 represent the human
XX polynucleotide sequences of the invention which are differentially
XX expressed during foam cell differentiation
XX
SQ Sequence 2889 BP; 600 A; 873 C; 929 G; 487 T; 0 U; 0 Other;
Query Match 89.7%; Score 2023.4; DB 6; Length 2889;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 2255; Conservative 0; Mismatches 1; Indels 211; Gaps 2;
QY 1 ATGGGCTTCTCTCTGAGCTGTGACGCCCCGAGGCGACGCGGGTCTCTGACGAATGACG 60
DB 188 ATGGGCTTCTCTCTGAGCTGTGACGCCCCGAGGCGACGCGGGTCTCTGACGAATGACG 247
QY 61 GAGGCCGAGCTTCTGTCTACTGAGGCGATGAGAAATGATGAGCCGCGGTCAGAGT 120
DB 248 GAGGCCGAGCTTCTGTCTACTGAGGCGATGAGAAATGATGAGCCGCGGTCAGAGT 307
QY 121 GACAGGAGATGATGACAGACTGCTTCAACATGTCCTCTGACGACAGTGGGGCCGAGAC 180
DB 308 GACAGGAGATGATGACAGACTGCTTCAACATGTCCTCTGACGACAGTGGGGCCGAGAC 367
QY 181 CGGGCCATCAAGCCCTGACAGCCCCCATCAATCAATCTCTGGGCTGAGATCAACGCAAACT 240
DB 368 CGGGCCATCAAGCCCTGACAGCCCCCATCAATCAATCTCTGGGCTGAGATCAACGCAAACT 427
QY 241 GA-GGGCTTGAGCCGCTTGTCTGCGGACGACGACGAGATCTGAATCAAGGCCCCCTGAG 299
DB 428 GAGGGCCCTGAGCCGCTTGTCTGCGGACGACGACGAGATCTGAATCAAGGCCCCCTGAG 487
QY 300 CAAGCTGAGCTTCTCTATCCGGGAGAGGCGACGCTTTCGCAACCTACAGGACGATG 359
DB 488 CAAGCTGAGCTTCTCTATCCGGGAGAGGCGACGCTTTCGCAACCTACAGGACGATG 547
QY 360 GCAGCAGCTGACAGAGAGCTCAACAAGCCACAGCAGAGCAATTGAGAAAGCTGAAAG 419
DB 548 GCAGCAGCTGACAGAGAGCTCAACAAGCCACAGCAGAGCAATTGAGAAAGCTGAAAG 607
QY 420 CCAAGTACCAAGCTCTGGCAGGAGCAGTGGCCCAAGCCCAAGCCAGTACCAAGAGCCAG 479
DB 608 CCAAGTACCAAGCTCTGGCAGGAGCAGTGGCCCAAGCCCAAGCCAGTACCAAGAGCCAG 667
QY 480 CAAAGACAAAGAACCTGTGACAAAGCCAAAGACAAAGTATGTGCGACGCTGTGAAAGCTTT 539
DB 668 CAAAGACAAAGAACCTGTGACAAAGCCAAAGACAAAGTATGTGCGACGCTGTGAAAGCTTT 727

QY 540 TGTCTACCAACACCCGCTATGTCTGCGGCTGCGGCTGCGACGTAACACCAAGACCA 599
DB 728 TGTCTACCAACACCCGCTATGTCTGCGGCTGCGGCTGCGACGTAACACCAAGACCA 787
QY 600 CCAACGACTCTCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
DB 788 CCAACGACTCTCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 847
QY 660 TTGCATCTCGAAGAGAGATCTCTGACAGAAATACCTGAGATTTAGAGCTGTGAGAGTA 719
DB 848 TTGCATCTCGAAGAGAGATCTCTGACAGAAATACCTGAGATTTAGAGCTGTGAGAGTA 907
QY 720 GGTGTGAGCAATTCACCGGAGATGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 779
DB 908 GGTGTGAGCAATTCACCGGAGATGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 967
QY 780 GTACCAAGGCTTCTGCGACAGTATGAGTCCGACCTGACGTCACCTGTCTACGTT 839
DB 968 GTACCAAGGCTTCTGCGACAGTATGAGTCCGACCTGACGTCACCTGTCTACGTT 1027
QY 840 CGATGATGACTGCTTGAAGAGAGTGAACCGCTGAGCCTGAGGAGCTCAGCTGAACGA 899
DB 1028 CGATGATGACTGCTTGAAGAGAGTGAACCGCTGAGCCTGAGGAGCTCAGCTGAACGA 1087
QY 900 GCTGACTGTGAGAGCGTGCACACACGCTGACCTCAGTGAACAGATGAGCTGTGCG 959
DB 1088 GCTGACTGTGAGAGCGTGCACACACGCTGACCTCAGTGAACAGATGAGCTGTGCG 1147
QY 960 CACCGAGATGCTGCTTCAAGCGGCGAGAGATGCTTACGACGCTGCAACAGAGTCCGGA 1019
DB 1148 CACCGAGATGCTTCAAGCGGCGAGAGATGCTTACGACGCTGCAACAGAGTCCGGA 1207
QY 1020 TGAAGAGAGAAACCCACCCCGGAGGAGGAGTCACTGCTGAGGAGAGGCAATGCT 1079
DB 1208 TGAAGAGAGAAACCCACCCCGGAGGAGGAGTCACTGCTGAGGAGAGGCAATGCT 1267
QY 1080 GCAAGAGACCTGCAAGGAGCTGCAAGTACGCTGTGACCCAGGCCAAGCTGAGGCCCA 1139
DB 1268 GCAAGAGACCTGCAAGGAGCTGCAAGTACGCTGTGACCCAGGCCAAGCTGAGGCCCA 1327
QY 1140 GCAAGAGCTTCTGCAAGCACAAGTGTGAGCACTGGGCCCCCGGAGGCCCCGCTGTGCT 1199
DB 1328 GCAAGAGCTTCTGCAAGCACAAGTGTGAGCACTGGGCCCCCGGAGGCCCCGCTGTGCT 1387
QY 1200 GCTCTGAGAGATGACCGGCACTCCAGCTGCTCTGAGAGCAGAGGAGGAGGAG 1259
DB 1388 GCTCTGAGAGATGACCGGCACTCCAGCTGCTCTGAGAGCAGAGGAGGAGGAGGAG 1447
QY 1260 GACACCCAGCTGAGAGATCTTAAAGACCAATCTTCAAGAAATCTTCCGCCCAAGTTCTC 1319
DB 1448 GACACCCAGCTGAGAGATCTTAAAGACCAATCTTCAAGAAATCTTCCGCCCAAGTTCTC 1507
QY 1320 G----- 1320
DB 1508 GCTCTCTCAACCGCTGACGCTATTCGAGAGGTGCAAGGCCCTTCATGACGAGCTGTG 1567
QY 1321 ----- 1320
DB 1568 GTACCAAGGAGCCATCCGAGGAGAGAGTGGCTGAGCTGTGCTGCTGCTGCTGCTGCTG 1627
QY 1321 ----- 1320
DB 1628 CCTGTGCGGAGAGCCAGAGGCAAGAGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1687
QY 1321 -----AACTGTACCACTGAGAGAGGAGAGGCTT 1349
DB 1688 GCCCCGAGCACTTATCATCATCATCTTGTGATTAACCTGTACCGACTGAGAGGAGAGCTT 1747
QY 1350 TCTTATCATCTCTTGTGATGAGCACTTATGAGCACTTATGAGCACTTATGAGCACTTATGAG 1409
DB 1748 TCTTATCATCTCTTGTGATGAGCACTTATGAGCACTTATGAGCACTTATGAGCACTTATGAG 1807

QY 1410 GAGTGGTGTTCCTGCAACAGGCGGTGTCGCCAAGACAAAGTGGGTGTGAACCATGAGGA 1469
 DB 1808 GAGTGGTGTTCCTGCAACAGGCGGTGTCGCCAAGACAAAGTGGGTGTGAACCATGAGGA 1867
 QY 1470 CCTGTGTGTGGTGAACAGATTGAGCGGGGAACTTTGGCGAAGTGTTCAGCGGAGCGCTT 1529
 DB 1868 CCTGTGTGTGGTGAACAGATTGAGCGGGGAACTTTGGCGAAGTGTTCAGCGGAGCGCTT 1927
 QY 1530 GCGAGCGCAACACACCTCTGTGGCGGTGAAGTCTTTCGAGAGACGCTCCCACTTACCT 1589
 DB 1928 GCGAGCGCAACACACCTCTGTGGCGGTGAAGTCTTTCGAGAGACGCTCCCACTTACCT 1987
 QY 1590 CAAGGCGCAAGTTCTACAGGAAGCGAAGATCCCTGAAGCGATACAGCACCACCAACCTCG 1649
 DB 1988 CAAGGCGCAAGTTCTACAGGAAGCGAAGATCCCTGAAGCGATACAGCACCACCAACCTCG 2047
 QY 1650 GCGTCTCATTTGATGTGTGCAACCGAAGAGCCCATCTACATCGTCATGAGAGCTTGTGCA 1709
 DB 2048 GCGTCTCATTTGATGTGTGCAACCGAAGAGCCCATCTACATCGTCATGAGAGCTTGTGCA 2107
 QY 1710 GGGGGGCGACTTCTCTGACCTTCTCCGCAACGAGGGGGGCCGCTCGGGGTGAAGACTCT 1769
 DB 2108 GGGGGGCGACTTCTCTGACCTTCTCCGCAACGAGGGGGGCCGCTCGGGGTGAAGACTCT 2167
 QY 1770 GCTGCGAGATGTTGGGGGATGCAAGTCTGTGCAATGAGTACCTGAGAGCAAGTGTGCAT 1829
 DB 2168 GCTGCGAGATGTTGGGGGATGCAAGTCTGTGCAATGAGTACCTGAGAGCAAGTGTGCAT 2227
 QY 1830 CCAACCGGACCTTGGCTGTGCAAGTCTGTGCAAGTACCTGAGAGCAAGTGTGCAT 1889
 DB 2228 CCAACCGGACCTTGGCTGTGCAAGTCTGTGCAAGTACCTGAGAGCAAGTGTGCAT 2287
 QY 1890 TGACTTTGGAGTGTCCCGAGAGGAGCCGATGAGGAGTCTTATGACGCTCAGGGGCGCTCAG 1949
 DB 2288 TGACTTTGGAGTGTCCCGAGAGGAGCCGATGAGGAGTCTTATGACGCTCAGGGGCGCTCAG 2347
 QY 1950 ACAAGTCCCCCGTGAATGAGACCGCACCTGAGGCGCTTAACTAGGCGCGCTACTCTCCGA 2009
 DB 2348 ACAAGTCCCCCGTGAATGAGACCGCACCTGAGGCGCTTAACTAGGCGCGCTACTCTCCGA 2407
 QY 2010 AAGCGACGTGTGAGAGCTTTGGCACTTTGCTGTGAGAGACCTTACGCTGGGGGCGCTCCCG 2069
 DB 2408 AAGCGACGTGTGAGAGCTTTGGCACTTTGCTGTGAGAGACCTTACGCTGGGGGCGCTCCCG 2467
 QY 2070 CTATCCCAACCTCAGCAATCAGACACACGAGAGTTTGTGAGAGAGGGGGCGCTGTGCC 2129
 DB 2468 CTATCCCAACCTCAGCAATCAGACACACGAGAGTTTGTGAGAGAGGGGGCGCTGTGCC 2527
 QY 2130 CTGCCAGAGCTGTCTCGATGCGGTTCAGGCTTCATGAGACAGTGTGGGCGCTATGA 2189
 DB 2528 CTGCCAGAGCTGTCTCGATGCGGTTCAGGCTTCATGAGACAGTGTGGGCGCTATGA 2587
 QY 2190 GCGTGGGAGAGGCGGCGCAGCTTTCAGACACATCTACAGAGAGCTCAGAGACATCCGAAAGCG 2249
 DB 2588 GCGTGGGAGAGGCGGCGCAGCTTTCAGACACATCTACAGAGAGCTCAGAGACATCCGAAAGCG 2647
 QY 2250 GCATCGG 2256
 DB 2648 GCATCGG 2654
 RESULT 6
 ADR39816 standard; cDNA; 2623 BP.
 ADR39816;
 18-NOV-2004 (first entry)
 Human kinase and phosphatase KPP-43 encoding cDNA SEQ ID NO:89.
 human: kinase and phosphatase protein; KPP: enzyme; cytosolic;
 antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective;

KW cerebroprotective; anti-HIV; antiallergic; anti-inflammatory;
 KW thymimetic; gene therapy; cell proliferative disorder; cancer;
 KW atherosclerosis; neurological disorder; epilepsy; Huntington's disease;
 KW stroke; immune disorder; inflammatory disorder; AIDS; allergy;
 KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
 KW KPP-43; single nucleotide polymorphism; SNP; gene; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 95..2389
 FT /tag= a
 FT /product= "kinase and phosphatase KPP-43"
 FT variation replace(109,C)
 FT /tag= b
 FT /standard_name= "single nucleotide polymorphism (SNP)"
 PD MO2004074453-AZ.
 PD 02-SEP-2004.
 XX 20-FEB-2004; 2004MO-US005092.
 XX 20-FEB-2003; 2003US-0449059P.
 PR 19-MAR-2003; 2003US-0456932P.
 PR 28-MAR-2003; 2003US-0458844P.
 PR 09-APR-2003; 2003US-0461678P.
 PR 17-APR-2003; 2003US-0463937P.
 XX (INCY-) INCYTE CORP.
 PA Ramkumar J, Margulis JP, Swarnakar A, Chawla NK, Tran UK;
 PI Becha SD, Lee SY, Hafalia AJA, Richardson TW, Khare R, Jiang X;
 PI Jackson AA, Yang J, Gorvad AE;
 XX WPI: 2004-635568/61.
 DR P-PADB; ADR39770.
 XX New human kinases and phosphatases (KPP) for diagnosing, treating and
 PT preventing diseases or conditions associated with aberrant KPP expression
 PT e.g. cancer, acquired immunodeficiency syndrome, epilepsy, or infections.
 XX
 PS Claim 5; SEQ ID NO 89; 299pp; English.
 XX
 CC The present sequence encodes the human kinase and phosphatase protein
 CC (KPP), designated KPP-43. The human KPP sequences from the present
 CC invention have cytosolic, antiarteriosclerotic, anticonvulsant,
 CC nootropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,
 CC antiinflammatory and thymimetic activities, and can be used in gene
 CC therapy. The human KPP proteins and polynucleotides can be used in
 CC diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of KPP, such as cell
 CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome).
 CC disorders, or infections. They can also be used in assessing the effects
 CC of exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of KPP. The KPP or its fragments are useful in screening
 CC compounds for effectiveness as agonist or antagonist of the polypeptides,
 CC or in altering the expression of the target polynucleotide and compounds
 CC that specifically bind to or modulate the activity of the polypeptide.
 XX
 SQ Sequence 2623 BP; 541 A; 786 C; 840 G; 456 T; 0 U; 0 Other;
 XX
 XX Query Match 74.1%; Score 1671.6; DB 13; Length 2623;
 XX Best Local Similarity 84.3%; Pred. No. 1,1e-305;
 XX Matches 2078; Conservative 0; Mismatches 4; Indels 384; Gaps 2;
 QY 1 ATGGGCTTCTCTTCTGAGCTGTGACGCCCCCAAGGCGGCTCTGACGCAATGACG 60
 DB 95 ATGGGCTTCTCTTCTGAGCTGTGACGCCCCCAAGGCGGCTCTGACGCAATGACG 154
 QY 61 GAGGCGGAGCTTGTCTACTGAGGCGCATGAGAAATGAGATGGCCACGCGGTCAAGAGT 120

155 GAGCGGAGCTTGTCTACTGAGGCGATGAGAAATGATGCGCCAGCGGCTCAAGAGT 214
121 GACAGGAGATATGACAGAGCTGTTCAACATATCTCCCTGACAGACAGTGGGGCCAGAC 180
215 GACAGGAGATATGACAGAGCTGTTCAACATATCTCCCTGACAGAGAGTGGGGCCAGAC 274
181 CGGGCCATGACGCGCTTACAGAGCCCATCAATCAATCTCGGCTGAGATCAACAGCCAACT 240
275 CGGGCCATGACGCGCTTACAGAGCCCATCAATCAATCTCGGCTGAGATCAACAGCCAACT 305
241 GAGGCGCTGAGCGCTTGTCTGCGGACAGACAGAGAGATCTGAATCAAGGCGCCCTGAGC 300
306 ----- 305
301 AAGCTAGCTGCTCATCCGGGAAACGGACAGAGCTTCCGAAGACTTAAGCCAGAGACTGG 360
306 ----- 305
361 CAGCAGCTGACAGAGAGCTCAACCAAGACCAAGCAGCAGGACATTTGAGAGCTGAAGAGC 420
306 -----AGACCCACAGCCAGGACATTTGAGAGAGCTGAAGAGC 340
421 CAGTACCGAGCTTGGCAACGGGACAGTGCACCAAGCCAAAGCAAGTACAGAGAGCCAGC 480
341 CAGTACCGAGCTTGGCAACGGGACAGTGCACCAAGCCAAAGCAAGTACAGAGAGCCAGC 400
481 AAGACAAAGGACCGGTGACAAAGGCGCAAGAGAAATATGTGGCAGCCTGTGGAGAGCTTT 540
401 AAGACAAAGGACCGGTGACAAAGGCGCAAGAGAAATATGTGGCAGCCTGTGGAGAGCTTT 460
541 GCTCACCAACACGCTATGTGCTGGGCGTGCAGGCGCTGCGAGCTACACCAACCAAC 600
461 GCTCACCAACACGCTATGTGCTGGGCGTGCAGGCGCTGCGAGCTACACCAACCAAC 520
601 CACCAAGCTTGTGCTGCGCGGCTGCTGCGGTCACTGACAGAGACCTGCAACGAGAGATGCT 660
521 CACCAAGCTTGTGCTGCGCGGCTGCTGCGGTCACTGACAGAGACCTGCAACGAGAGATGCT 580
661 TGCATCTGAAGAGATCTCTGACAGAAATACCTGGAATTTAGCAGCCTGTGGCAGAGTAG 720
581 TGCATCTGAAGAGATCTCTGACAGAAATACCTGGAATTTAGCAGCCTGTGGCAGAGTAG 640
721 GTGGTGGCCATTACCGGAGAGATGAGCTGACAGTGCACCTGCAATCCAGCTGAGGCTGAG 780
641 GTGGTGGCCATTACCGGAGAGATGAGCTGACAGTGCACCTGCAATCCAGCTGAGGCTGAG 700
781 TACCAAGGCTTCTGCGACAGTATGGGTCCGCACTGACGTCCCACTGTGTCAAGTTTC 840
701 TACCAAGGCTTCTGCGACAGTATGGGTCCGCACTGACGTCCCACTGTGTCAAGTTTC 760
841 GATGAGCTCATGCTTGAAGAGAGGAGTAAACGCTGAGAGCTGGGGAGCTCCAGCTGAAGAG 900
761 GATGAGCTCATGCTTGAAGAGAGGAGTAAACGCTGAGAGCTGGGGAGCTCCAGCTGAAGAG 820
901 CTGACTGTGAGAGAGCTGACAGACAGCTGACCTCACTGATGACAGATGAGCTGGCTGTGGCC 960
821 CTGACTGTGAGAGAGCTGACAGACAGCTGACCTCACTGATGACAGATGAGCTGGCTGTGGCC 880
961 ACCGAGATGGTGTTCAGGCGGACAGAGATGTTTACGAGCTGCAACAGAGCTTCCGGAAT 1020
881 ACCGAGATGGTGTTCAGGCGGACAGAGATGTTTACGAGCTGCAACAGAGCTTCCGGAAT 940
1021 GAAGAGAGAGAACCCCAACCCCGGAGAGAGGAGTCAAGTGTGGGCAAGAGAGAGAGTCTG 1080
941 GAAGAGAGAGAACCCCAACCCCGGAGAGAGGAGTCAAGTGTGGGCAAGAGAGAGAGTCTG 1000
1081 CAAGAGAGAGAGAGAGAGAGAGAGTGTGAGCTGTGACAGAGAGCAAGAGAGAGAGAGAG 1140
1001 CAAGAGAGAGAGAGAGAGAGAGAGTGTGAGCTGTGACAGAGAGAGAGAGAGAGAGAGAG 1060
1141 CAGAGAGTGTGAG 1200

1061 CAGAGATTGCTCAGACCAAGCTGAGAGACATGGGCGCGGAGAGCCCGCTGTGCTG 1120
1201 CTCCTGACAGATGACCGCACTTCAAGTGTCTCTCGAGACAGAGAGAGAGAGAGAGAGAG 1260
1121 CTCCTGACAGATGACCGCACTTCAAGTGTCTCTCGAGACAGAGAGAGAGAGAGAGAGAG 1180
1261 ACACCCACGCTGAGAGATCTTTAAGAGCCATCTCAGAAATTTCCGCCCAAGTTCTG 1320
1181 ACACCCACGCTGAGAGATCTTTAAGAGCCATCTCAGAAATCTTCCGCCCAAGTTCTG 1240
1321 ----- 1320
1241 CTCCTTCCACCGCTGACGCTCATTCGAGAGTGCAGAAAGCCCTGTGATGACAGCTGTG 1300
1321 ----- 1320
1301 TACACAGGAGCCATCCGAGGACAGAGTGGCTGAGCTGTGTGACATCTTGGGAGCTTC 1360
1321 ----- 1320
1361 CTGGTGGGAGAGAGCAGGAGCAAGCAGAGTACGTGTGTGCTGTGTGATGTCTG 1420
1321 -----AACCTGTACCGACTGGAAGGGAGAGCTTT 1350
1421 CCGCGCACTTGCATCATTCAGTCTTTGATTAACCTGTACCTGGAAGGGAGAGCTTT 1480
1351 CCTAGCATTCCTTGTGCTCATCGACCACTACTGAGACCCAGAGAGAGAGAGAGAGAGAG 1410
1481 CCTAGCATTCCTTGTGCTCATCGACCACTGTAGAGACCCAGAGAGAGAGAGAGAGAGAG 1540
1411 AGTGTGTGTCTCTGACAGAGGCTGTGCCAAGACAAAGTGGTGTGAACCATGAGAGAC 1470
1541 AGTGTGTGTCTCTGACAGAGGCTGTGCCAAGACAAAGTGGTGTGAACCATGAGAGAC 1600
1471 CTGGTGTGTGTGAGAGATTTGAGACGGGGGAACTTTGGGAAAGTGTTCAGAGAGAGAG 1530
1601 CTGGTGTGTGTGAGAGATTTGAGACGGGGGAACTTTGGGAAAGTGTTCAGAGAGAGAG 1660
1531 CGAGCCGACAAACCTGTGTGAGAGAGTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1590
1661 CGAGCCGACAAACCTGTGTGAGAGAGTCTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 1720
1591 AAGGCCAAGTTTCTACAGAGAGAGAGATCTGAGACAGTACAGCAACCCCAACATGTG 1650
1721 AAGGCCAAGTTTCTACAGAGAGAGAGATCTGAGACAGTACAGCAACCCCAACATGTG 1780
1651 CGTCTATGTGTGTCTGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1710
1781 CGTCTATGTGTGTCTGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1840
1711 GGGGCGGACCTTCTGACCTTCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1770
1841 GGGGCGGACCTTCTGACCTTCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900
1771 CTGACAGATGGTGGAGATGACAGCTGTGACATGAGATACCTGAGAGAGAGAGAGAGAG 1830
1901 CTGACAGATGGTGGAGATGACAGCTGTGACATGAGATACCTGAGAGAGAGAGAGAGAG 1960
1831 CACCGGAGACCTGCTGTGAGAACTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1890
1961 CACCGGAGACCTGCTGTGAGAACTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2020
1891 GACTTGGAGATGCTCGAG 1950
2021 GACTTGGAGATGCTCGAG 2080
1951 CAAATCCCGTGAAGTGAACCGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2010
2081 CAAATCCCGTGAAGTGAACCGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2140
2011 AGCAGAGTGTGAG 2070
2141 AGCAGAGTGTGAG 2200

QY 2071 TATCCAACTCAGCATATGAGACAGACGGGAGTTTGTGAGAGGGGGCCGCTGCCC 2130
DB 2201 TATCCAACTCAGCATATGAGACAGACGGGAGTTTGTGAGAGGGGGCCGCTGCCC 2260
QY 2131 TGCACAGCTGTGTCTGTATGCGCTGTTCAGAGCTATGAGAGCTGTGAGCTATGAG 2190
DB 2261 TGCACAGCTGTGTCTGTATGCGCTGTTCAGAGCTATGAGAGCTGTGAGCTATGAG 2320
QY 2191 CCTGGGACGGGCGCCAGCTTTCAGACCATCTTACAGAGAGCTGACAGCATCCGAAGCCG 2250
DB 2321 CCTGGGACGGGCGCCAGCTTTCAGACCATCTTACAGAGAGCTGACAGCATCCGAAGCCG 2380
QY 2251 CATCGG 2256
DB 2381 CATCGG 2386

RESULT 7

ADH43089

ID ADH43089 standard; cDNA; 2680 BP.

XX ADH43089;

XX 25-MAR-2004 (first entry)

XX CRAM protein related cDNA #SEQ ID 1.

XX Neuroprotective; neutralisation; CRAM;

XX collapsing-response mediator protein-associated molecule; gene therapy;

XX mitochondria; drug development; neural disease; pathosis; gene; ss.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS 60..2522

XX /*tag= a

XX WO2004001038-A1.

XX 31-DEC-2003.

XX 19-JUN-2003; 2003WO-JP007766.

XX 19-JUN-2002; 2002JP-00179105.

XX (NEW!) NEW IND RES ORG.

XX Yanagi S;

XX WPI; 2004-099123/10.

XX P-PSDB; ADH43090.

XX Gene and proteins participating in neutralization of cells or tissues,

XX useful in gene therapy and regeneration medicine, applicable in

XX diagnosis, drug development for neural diseases and study of mechanism of

XX pathosis.

XX Claim 5; SEQ ID NO 1; 101bp; Japanese.

XX The invention relates to a method for inducing neutralisation of cells or

XX tissues by using a protein binding to CRAM (collapsing-response mediator

XX protein-associated molecule) protein or its encoded gene. The proteins

XX and their encoded genes are useful in gene therapy and regenerative

XX medicine, e.g. by inducing neutralisation of mitochondria. They are also

XX applicable in diagnosis, drug development for neural diseases and

XX studying the mechanism of pathosis. The current sequence represents CRAM

XX protein related cDNA.

XX Sequence 2680 BP; 629 A; 758 C; 791 G; 502 T; 0 U; 0 Other;

XX Query Match 68.3%; Score 1540.4; DB 12; Length 2680;

XX Best Local Similarity 79.2%; Pred. No. 5.9e-281;

Matches 1954; Conservative 0; Mismatches 296; Indels 216; Gaps 2;

QY 1 ATGGGCTTCTCTTCTGAGCTGTGCAAGCCCCCAAGGGTCTTGCAGCAAAATGCAAG 60
DB 60 ATGGGCTTCTCTTCTGAGCTGTGCAAGCCCCCAAGGGTCTTGCAGCAAAATGCAAG 119
QY 61 GAGGCGAGCTTGTCTACTAGAGGGCATGAGAAATGATGGGCCAGGGGCTCAAGAGT 120
DB 120 GAGGCGAGCTTGTCTACTAGAGGGCATGAGAAATGATGGGCCAGGGGCTCAAGAGT 179
QY 121 GACAGGAGTATGACAGAGCTGTTCACCAATGCTCTCTGACAGAGAGTGGGGCCAGAG 180
DB 180 GACAGGAGTATGACAGAGCTGTTCACCAATGCTCTCTGACAGAGAGTGGGGCCAGAG 229
QY 181 CGGAGCATCAGCCCTGACAGCCCTCATAGTCAAGTCTTGGGCTGAGATCAGCCAAAT 240
DB 240 TGGAGCAGGCGCCCGACAGCCCTGACAGCCCTCATAGTCAAGTCTTGGGCTGAGATC 299
QY 241 GAGGCGCTGAGCGGCTGTGCGGGGACAGAGAGATCTGAATCAGGGCCCTGAGC 300
DB 300 GAGAGCTTGAAGCGGCTGTGCGGGGACAGAGATCTGAATCAGGGCCCTGAGC 359
QY 301 AAGCTGAGCTGTCTCATCCGGGAACGGGACAGAGCTTGCAGAGCTTACAGAGAGTGG 360
DB 360 AAGCTGAGCTGTCTCATCCGGGAACGGGACAGAGCTTGCAGAGCTTACAGAGAGTGG 419
QY 361 CAGCAGCTGACAGAGAGCTTACCAAGACCCACAGCAGAGCATTTGAGAAAGTGAAGAC 420
DB 420 CAGCAGCTGACAGAGAGCTTACCAAGACCCACAGCAGAGCATTTGAGAAAGTGAAGAC 479
QY 421 CAGTACCGAGCTGTGCAAGGACAGTGGCCCAAGCCCAAGCAAGTACAGAGAGGCGAG 480
DB 480 CAGTACCGAGCTGTGCAAGGACAGTGGCCCAAGCCCAAGCAAGTACAGAGAGGCGAG 539
QY 481 AAAGCAAGAGCTGTGCAAGGACAGGACCAAGATATGTGCGAGCTGTGAGAGCTTT 540
DB 540 AAAGCAAGAGCTGTGCAAGGACAGGACCAAGATATGTGCGAGCTGTGAGAGCTTT 599
QY 541 GCTCACCACACCGCTATGTGCTGGGCGTGGCGGCTGCGAGTACACCAAGCACAC 600
DB 600 GCTCACCACACCGCTATGTGCTGGGCGTGGCGGCTGCGAGTACACCAAGCACAC 659
QY 601 CACGAGCTGTGCTGCGCGGCGCTGCTGCGGCTGCTGCAAGAGCTTGAAGAGTGGCT 660
DB 660 CACGAGCTGTGCTGCGCGGCGCTGCTGCGGCTGCTGCAAGAGCTTGAAGAGTGGCT 719
QY 661 TGCATCTGAAGAGATCTGAGAGATCTGAGAGATTTAGAGAGCTGTGAGAGTGA 720
DB 720 TGCATCTGAAGAGATCTGAGAGATCTGAGAGATTTAGAGAGCTGTGAGAGTGA 779
QY 721 GTGTGGCCATTACCGGAGATGCTGAGCTGTGCGGATTCAGGCTTGAAGCTGAG 780
DB 780 GTGTGGCCATTACCGGAGATGCTGAGCTGTGCGGATTCAGGCTTGAAGCTGAG 839
QY 781 TACCAAGGCTTCTGAGAGATGATGCTGCGAGCTGAGAGTCCACCCGTGTGAGCTTC 840
DB 840 TACCAAGGCTTCTGAGAGATGATGCTGCGAGCTGAGAGTCCACCCGTGTGAGCTTC 899
QY 841 GATGAGTCACTGCTTGAAGAGGCTGAGAGCTGAGAGCTGAGAGCTTGAAGCTGAG 900
DB 900 GATGAGTCACTGCTTGAAGAGGCTGAGAGCTGAGAGCTGAGAGCTTGAAGCTGAG 959
QY 901 GTGATCTGAGAGAGCTGAGAGCAACGCTGAGCTTCAAGTACAGATGAGCTGTGAGCC 960
DB 960 GTGATCTGAGAGAGCTGAGAGCAACGCTGAGCTTCAAGTACAGATGAGCTGTGAGCC 1019
QY 961 ACCGAGATGCTTGAAGAGGCTGAGAGATGCTTGAAGAGCTGAGAGCTTGAAGAT 1020
DB 1020 ACCGAGATGCTTGAAGAGGCTGAGAGATGCTTGAAGAGCTGAGAGCTTGAAGAT 1079
QY 1021 GAAG 1080
DB 1080 GAAG 1139

CC on 27-AUG-2003 to correct OS field.)
XX Sequence 856 BP; 164 A; 265 C; 259 G; 168 T; 0 U; 0 Other;
SQ
Query Match 28.1%; Score 635; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 2.8e-110;
Matches 635; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1622 TGAAGCAGTACAGCCCAACATCTGCTCATTTGCTGTGACCCAGAGACG 1681
DB 1 TGAAGCAGTACAGCCCAACATCTGCTCATTTGCTGTGACCCAGAGACG 60
QY 1682 CCATCTACATCTGCTGTGAGAGCTTGTGAGGGGGGCGACTTCGACCTTCTCGACCG 1741
DB 61 CCATCTACATCTGCTGTGAGAGCTTGTGAGGGGGGCGACTTCGACCTTCTCGACCG 120
QY 1742 AGGGGGCCCGCTGCGGGGTGAAGACTCTGCTGAGATGTGTGGGGGATGACAGTGTGGCA 1801
DB 121 AGGGGGCCCGCTGCGGGGTGAAGACTCTGCTGAGATGTGTGGGGGATGACAGTGTGGCA 180
QY 1802 TGAAGTACCTGGAAGAGCACTGCTGCATCCACCGGACCTTGCTGCTCGAACTGCTCG 1861
DB 181 TGAAGTACCTGGAAGAGCACTGCTGCATCCACCGGACCTTGCTGCTCGAACTGCTCG 240
QY 1862 TGAAGAGAGAGATGCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATG 1921
DB 241 TGAAGAGAGAGATGCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATG 300
QY 1922 GGGTCTATGAGAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCGACCTGAGG 1981
DB 301 GGGTCTATGAGAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCGACCTGAGG 360
QY 1982 CCTTATCTATGAGAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCTTGTGCTCT 2041
DB 361 CCTTATCTATGAGAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCTTGTGCTCT 420
QY 2042 GGGAGACCTTACAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCTTGTGCTCT 2101
DB 421 GGGAGACCTTACAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCTTGTGCTCT 480
QY 2102 AGTTTGTGAGAGAGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2161
DB 481 AGTTTGTGAGAGAGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 2162 GGGTCTATGAGAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCTTGTGCTCT 2221
DB 541 GGGTCTATGAGAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCTTGTGCTCT 600
QY 2222 ACCAGAGCTGAGAGAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCTTGTGCTCT 2256
DB 601 ACCAGAGCTGAGAGAGCTGAGGGGGCTGAGCAAGTCCCTGTAAGTGGACCTTGTGCTCT 635
RESULT 9
ADL71052
ID ADL71052 standard; DNA; 2818 BP.
XX
XX ADL71052;
DT 20-MAY-2004 (first entry)
XX
XX Gene encoding type II collagen expression promoting protein, seq id 45.
DE
XX Osteopathic; antiinflammatory; antirheumatic; antiarthritic;
XX gene therapy; type II collagen; expression; cartilage disease;
KM osteoarthritis; cartilage defect; rheumatoid arthritis; human; gene; ds.
XX
OS Homo sapiens.
XX
XX MO2003087375-A1.
XX
XX 23-OCT-2003.
XX

PF 16-APR-2003; 2003MO-JP004802.
XX
XX 16-APR-2002; 2002JP-00113908.
PR 19-APR-2002; 2002US-0373594P.
XX
XX (ASAH) ASAH KASEI KK.
PI Matsuda A, Honda G, Muramatsu S;
XX
XX WPI; 2003-845331/78.
DR P-PSDB; ADL71053.
XX
XX New purified protein that promotes type II collagen expression, useful
PT for preventing and treating a cartilage disease, e.g. osteoarthritis,
PT cartilage defect, or rheumatoid arthritis.
XX
XX Claim 4; SEQ ID NO 45; 271pp; English.
XX
XX The invention relates to a purified protein (I) that promotes type II
CC collagen expression. Also disclosed is an isolated polynucleotide
CC encoding (I), a recombinant vector comprising the polynucleotide and a
CC gene therapeutic agent comprising the recombinant vector as an active
CC ingredient. The proteins, genes, agents and methods are useful for
CC preventing and treating a cartilage disease, e.g. osteoarthritis,
CC cartilage defect, or rheumatoid arthritis. The current sequence
CC represents a human gene of the invention encoding a protein that promotes
XX type II collagen expression.
SQ Sequence 2818 BP; 909 A; 558 C; 684 G; 667 T; 0 U; 0 Other;
Query Match 16.6%; Score 375; DB 1; Length 2818;
Best Local Similarity 64.8%; Pred. No. 3.4e-61;
Matches 606; Conservative 0; Mismatches 520; Indels 9; Gaps 3;
QY 1316 TCTGCACTGTACACGAGTGAAGGGAGGCTTCTCTGACATTCCTTCTCATGAGACC 1375
DB 1686 TCGATATATCTGTATGATGATGAGGGACCGGGTTTCAAAATATCCCGAGTTATGATC 1745
QY 1376 ACCTACTGAGCACCCAGACCCCTGACCAAGAGAGTGTGTGCTTCTGACAGGGCTG 1435
DB 1746 ACCACTTCAATCAAGAGCAAGTATCATCAGAGAGTCTGGGGTGTGTGCTCAACCCCA 1805
QY 1436 TGGCCAAAGACAG---TGGGTCTGAACATGAGAGACCTGCTGTGGTGGTGAAGCATTC 1492
DB 1806 TCCCAAGAGATGAAGATGGTCTCAATCATGAATATGTTTCATTTGGAGAAATTA 1865
QY 1493 GACGGGGGAACCTTGGCGAAGTGTTCAGGAGAGCGCTGCGAGCCGACCAACACCTGATG 1552
DB 1866 GCAAGGGGAATTTTGTGTAAGTGTATAGG---GCACATTAAGATTAACCTCTGTTG 1922
QY 1553 CGGTGAAGTCTTGTGAGAGAGCGCTCCACCTGACCTCAAGGCCAATTTTCAAGAGAG 1612
DB 1923 CCATTAAGAGTGTGAGAGAGAGACCTTCTCAGAGATTAAGATTTTCTACAGAGAG 1982
QY 1613 CGAGATCTCTGAAGAGCTACAGCCCAACATGCTGCTGCTCATTTGCTGTGACCC 1672
DB 1983 CCAGAAATCTGAGAGCAATATGATCATCCCAATATGTCAGATGAGCGTGTGACAC 2042
QY 1673 AGAAGAGCCCATCTACATCTGATGATGATGATGATGATGATGATGATGATGATGATG 1732
DB 2043 AAGAGAGCTGTCTACATCATATTAATGAACTGTCCAGAGGGGTGATTTTCTACATTC 2102
QY 1733 TCCGACAGAGGGGGCCCGCTGCGGGGTGAAGACTCTGCTGAGATGTGTGGGGATG 1792
DB 2103 TGAAGAGAGAGAGAGAGAGCTGAGAGCTGAGAGAGAGTGTGAGATTTTCTTGAAGTTG 2162
QY 1793 CTGCTGGCATGAGATCTCTGAGAGAGAGAGTGTGATCCACCGGAGCTTGCTGCTGCA 1852
DB 2163 CTGCTGGCATGAGTGTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2222
QY 1853 ACTGCTGTGACAGAGAGAGATGTCTGAGAGTCAAGTCACTTGTGGATGTCCCGAGAG 1912
DB 2223 ACTGCTGTGAGAGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2282

XX 23-OCT-2003.
PD 16-APR-2003; 2003WO-JP004802.
XX 16-APR-2002; 2002JP-00113908.
XX 19-APR-2002; 2002US-0373594P.
XX (ASAH) ASAH KASEI KK.
XX Matsumura A, Honda G, Muramatsu S;
XX MPI: 2003-045331/78.
XX P-PSDB; ADL71055.
XX New purified protein that promotes type II collagen expression, useful
XX for preventing and treating a cartilage disease, e.g. osteoarthritis,
XX cartilage defect, or rheumatoid arthritis.
XX Claim 4; SEQ ID NO 47; 271bp; English.
XX The invention relates to a purified protein (I) that promotes type II
XX collagen expression. Also disclosed is an isolated polynucleotide
XX encoding (I), a recombinant vector comprising the polynucleotide and a
XX gene therapeutic agent comprising the recombinant vector as an active
XX ingredient. The proteins, genes, agents and methods are useful for
XX preventing and treating a cartilage disease, e.g. osteoarthritis,
XX cartilage defect, or rheumatoid arthritis. The current sequence
XX represents a human gene of the invention encoding a protein that promotes
XX type II collagen expression.
XX Sequence 2584 BP; 922 A; 446 C; 565 G; 651 T; 0 U; 0 Other:
XX
XX Query Match 14.6%; Score 328.8; DB 11; Length 2584;
XX Best Local Similarity 62.0%; Pred. No. 1.7e-52;
XX Matches 573; Conservative 0; Mismatches 342; Indels 9; Gaps 3;
XX
QY 1321 AACCTGACGATGGAAGGGAAGCTTCTTCTGATCTTCTGATGACCACTA 1380
DB 1552 AACATATATGATTCGAGGCGCTGAGTTTCAAACTTCTCACTTATGATCATC 1611
QY 1381 CTGAGCAGCCAGGAGCCCTCAGCAAGAGTGTGTCTCTGCAAGGCGCTGAGCC 1440
DB 1612 TATACAAACAAACAGCTCATCTAAGAAATCAGGTGATTTCTGATCTTATTTCT 1671
QY 1441 AAGGACAG--TGCGTCTGAACATGAGACCTGGTGTGAGAGCATTTGACCG 1497
DB 1672 AAGGACAGAAATGATTTCTCATGATGAAGATGATTTGGAGAAATTAAGTGGCAAG 1731
QY 1498 GGGAACTTTGGCAAGTGTTCAGCGAGCGCTGCGAGCCAGCAACCTGTGGCGGTG 1557
DB 1732 GGAATTTTGTGTAAGATATTAAG--GCACATTAAGATTAACCTTCTGTGTGTT 1788
QY 1558 AAGTCTTGTGAGAGACGCTCCACCTGACCTCAAGGCAAGTTTCTACAGAGAGAG 1617
DB 1789 AAAACATGTAAGAAATCTTCTCTCAGGAATTAATAATTTTACAGAGAGCAAA 1848
QY 1618 ATCTGTAAGCAGTACCAACCCCAATCTGCTCTCATTTGATGTCACACCAAG 1677
DB 1849 ATTCTCAGCAATATGATATCCCAATTTGTCAAACTTATAGAGTTTGCAACAAAGA 1908
QY 1678 CAGCCCATCTACATGCTCATGAGCTTTGAGAGGGGGGCACTTCTGACCTTCTCCG 1737
DB 1909 CAGCTGTCTACATCTTATGAACTGTGTTTCAGAGGTGATTTCTCAGCTTTGAGA 1968
QY 1738 AGGAGAGGGGGCGCGCGGGGTAAGACCTGCTGAGATGATGGGGGATGAGCTGCT 1797
DB 1969 AGGAGAGGATGATCAAAAATCAAACTAGTGAATTTTCAATTAGCGTGTGCT 2028
QY 1798 GGCATGAGTACCTGAGAGCAAGTGTGATCCACCGGAGCTGTGCTGCAAACTGC 1857
DB 2029 GGTATTTGTATCTCAGAGTAAATAAATCTGATACAGGGAAGCTTGTGCAAAATCTGC 2088

QY 1858 CTGTGACAGAGAAATGCTTCAAGATCAGTACTTTGGAGTCCGAGAGAGACC 1917
DB 2089 CTGTGATGTAATAATATGTTCTGAAATCAGTACTTTGAAATGCTGCTCAAGAGAT 2148
QY 1918 GATGGGCTCTATGACAGCTCAGGGGCTCAGACAGTCCCGTGAATGACCGGACCT 1977
DB 2149 GGTGAGTGTATTCATCTTCT--TGCTTAAGGCAATTCATTAATGAGACGACCG 2205
QY 1978 GAGGCCCTTAATACAGCCGCTTACTCTCCGAAAGCAGCTGTGAGCTTTGGCATCTTG 2037
DB 2206 GAAGCTTTAATATGAGATACAGTTCAAGAGAGTGAAGTGTGAGCTTTGGCATCTT 2265
QY 2038 CTCTGGAGACCTTCAAGCTCGGGGCGCTCCCTTATCCCACTGCAATACAGAGACA 2097
DB 2266 CTCTGGAGACCTTCAAGCTTCAAGGCTTGTCTGATCCGTAATGCAATATCAGAGCA 2325
QY 2098 CCGAGATTTGTGAGAGAGGGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2157
DB 2326 AAGAGCAAGTAAAGAGATACCGAGTGTACGCTCCCAAGCACTTCAAGAGATTT 2385
QY 2158 TTCAGGCTCATGAGAGCTGTGAGCTTATGAGCTTGGGCAAGGCGCCAGCTTACGACC 2217
DB 2386 TCCAAATCATGATGAAAGTGTGAGTATTAACCTGAAATCGCCCTAAGTTCACTGAA 2445
QY 2218 ATCTACAGAGCTGAGAGCATC 2241
DB 2446 CTTGAGAAAGAGCTCATCTATCATC 2469

RESULT 12

ACH14291
ID ACH14291 standard; cDNA; 449 BP.
XX ACH14291;
XX AC 13-OCT-2003 (first entry)
XX DT 13-OCT-2003 (first entry)
XX DE Human adult brain cDNA #1503.
XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX KW genome mapping; biodiversity; genetic disorder.
XX OS Homo sapiens.
XX PN US2003073623-A1.
XX PD 17-APR-2003.
XX PF 30-JUL-2001; 2001US-00918995.
XX PR 30-JUL-2001; 2001US-00918995.
XX PA (DRMA/) DRMANAC R T.
XX PA (LABA/) LABAT I.
XX PA (STAC/) STACHE-CRAIN B.
XX PA (DICK/) DICKSON M C.
XX PA (JONE/) JONES L W.
XX Demanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX MPI: 2003-615964/58.
XX DR New polynucleotide sequences obtained from various cDNA libraries, useful
XX PT as hybridization probes, as oligomers for PCR, for chromosome and gene
XX PT mapping, in the recombinant production of protein, or in generating
XX PT antisense DNA or RNA.
XX Claim 1; SEQ ID NO 1503; 449bp; English.
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to

CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostic assays as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome.
CC In forensic, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20030073623
CC
XX
SQ Sequence 449 BP; 90 A; 124 C; 148 G; 83 T; 0 U; 4 Other;

Query Match 14.5%; Score 326.8; DB 9; Length 449;
Best Local Similarity 99.4%; Pred. No. 3.3e-52;
Matches 328; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1707 GCAGGGGGGGGAGCTTCTGACCTTCCGCGCAGAGAGGGGGCCGCTGGGGTGAAGAC 1766
Db 120 GCAGGGGGGGGAGCTTCTGACCTTCCGCGCAGAGAGGGGGCCGCTGGGGTGAAGAC 179
QY 1767 TCTGCTGAGATGATGGGGGATCAGCTGCTGAGATGAGATCCTGAGAGCAAGTCTG 1826
Db 180 TCTGCTGAGATGATGGGGGATCAGCTGCTGAGATGAGATCCTGAGAGCAAGTCTG 239
QY 1827 CATCCACCGGAGCTGCTGCTGAGATGCTGCTGAGATGAGATGCTGAGATG 1886
Db 240 CATCCACCGGAGCTGCTGCTGAGATGCTGCTGAGATGAGATGCTGAGATG 299
QY 1887 CATGCTTTGGGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1946
Db 300 CATGCTTTGGGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
QY 1947 CAGACAGATCCCGGTGAGATGAGACCGACCTGAGCCCTTAAGTACGGCCGCTACTCTC 2006
Db 360 CAGACAGATCCCGGTGAGATGAGACCGACCTGAGCCCTTAAGTACGGCCGCTACTCTC 419
QY 2007 CGAAGCGAGCTGTGAGCTTTGGCATCTT 2036
Db 420 CGAAGCGAGCTGTGAGCTTTGGCATCTT 449

RESULT 13
AAS17052
ID AAS17052 standard; cDNA; 361 BP.
XX
XX AAS17052;
XX
DT 14-FEB-2002 (first entry)
XX
DE Human protein kinase cDNA mine16676human_g1.
XX
XX Human; 88; protein kinase; mine16676human_g1; cytosolic; antiangiinal;
XX hypotensive; cardiant; cardiovascular disorder; heart failure;
XX hypertension; atrial fibrillation; dilated cardiomyopathy;
XX idiopathic cardiomyopathy; angina; proliferative disorder; cancer;
XX melanoma; prostate cancer; cervical cancer; breast cancer; colon sarcoma;
XX gene therapy.
XX
OS Homo sapiens.
XX
XX US6309849-B1.
XX
XX 30-OCT-2001.
XX
XX 31-AUG-1999; 99US-00387212.
XX
XX 31-AUG-1999; 99US-00387212.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
PA

XX
PI Robison KE;
XX
XX WPI; 2002-048371/06.
DR
XX
XX Identifying compound which binds to a kinase, useful for treating
XX diseases e.g. cancer, by contacting kinase with test compound and
XX detecting its binding to the kinase.
XX
PS Claim 1; Fig 5; 45pp; English.
XX
XX
XX The invention relates to identifying a compound which binds to a kinase
XX (encoded by a nucleotide sequence of 1868, 403, 545, 361, 473, 3001, 526,
XX 683 or 1448 base pair (bp) as given in the specification) comprising
XX contacting the kinase with a test compound under suitable conditions for
XX binding, and detecting binding of the compound to the kinase. The method
XX is useful for identifying a compound which binds to the kinase and also
XX for isolating compounds which modify the activity of the kinase. The
XX identified compounds are useful for treating a subject having a disorder
XX characterised by aberrant kinase activity where the disorder includes
XX cellular growth related disorders which includes a disorder, disease, or
XX condition characterised by a deregulation, e.g. an upregulation or a
XX downregulation, of cellular growth, cellular growth deregulation due to
XX deregulation of cellular proliferation, cell cycle progression, cellular
XX differentiation and/or cellular hypertrophy, cardiovascular disorders
XX such as heart failure, hypertension, atrial fibrillation, dilated
XX cardiomyopathy, idiopathic cardiomyopathy, or anginal, proliferative
XX disorders such as cancer (including melanoma, prostate cancer, cervical,
XX breast, colon sarcoma). The kinases and antibodies raised against them
XX are useful in one or more method such as screening assays, predictive
XX medicine and methods of treatment. The nucleic acid molecules are useful
XX for expressing kinase and phosphatase protein (e.g. in gene therapy
XX applications), to detect kinase and phosphatase mRNA or a genetic
XX alteration in a kinase and phosphatase gene and to modulate kinase and
XX phosphatase activity. The present sequence is a cDNA for a human protein
XX kinase
XX
SQ Sequence 361 BP; 71 A; 101 C; 120 G; 68 T; 0 U; 1 Other;

Query Match 13.1%; Score 296.2; DB 6; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-46;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;

QY 1613 CGAGATCTCTGAGAGCAGTACAGACCCCAACATGCTGCTGATGCTGACACC 1672
Db 20 CNAAGATCTCTGAGAGCAGTACAGACCCCAACATGCTGCTGATGCTGACACC 79
QY 1673 AGAAGCAGCCCATCTACATGCTGATGAGCTTGTGAGAGGGGGGCACTTCTGACCTTCC 1732
Db 80 AGAAGCAGCCCATCTACATGCTGATGAGCTTGTGAGAGGGGGGCACTTCTGACCTTCC 139
QY 1733 TCCGACCGAGGGGGGGCCGCTGCGGGGTGAAGACTGCTGAGATGAGTGGGGATGACG 1792
Db 140 TCCGACCGAGGGGGGGCCGCTGCGGGGTGAAGACTGCTGAGATGAGTGGGGATGACG 199
QY 1793 CTGCTGAGTGAAGTACTGAGAGAGCAAGTGTGATCCACCGGAGACTGTGCTGCGA 1852
Db 200 CTGCTGAGTGAAGTACTGAGAGAGCAAGTGTGATCCACCGGAGACTGTGCTGCGA 259
QY 1853 ACTGCTGTGAGAGAGAGATGCTGAGAG-ATCAGTACCTTT-GGATGTGCTGAGAG 1910
Db 260 ACTGCTGTGAGAGAGAGATGCTGAGAGATGAGATGAGATGCTGAGAGATGCTGAGAG 319
QY 1911 GGAAGCC---GATGGGGTCTATGACAGCTCAGGGGGGCTCAG 1949
Db 320 GGAAGCCGATGGGGGTCTATGACAGCTCAGGGGGGCTCAG 361

RESULT 14
AAD61604
ID AAD61604 standard; cDNA; 361 BP.
XX
XX AAD61604;
XX
AC

```
XX
DT 15-JUN-2004 (first entry)
XX
DE Human protein kinase CDNA, 16676S1.
XX
KW Kinase; phosphatase; drug target; therapy; gene; human; ss.
XX
OS Homo sapiens.
XX
PN US2003104505-A1.
XX
PD 05-JUN-2003.
XX
PF 12-APR-2002; 2002US-00121925.
XX
PR 31-AUG-1999; 99US-00387212.
XX
PR 07-SEP-2001; 2001US-00948802.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Robison KE;
XX
DR WPI; 2003-801234/75.
XX
PT Detecting and modulating the activity of specified kinases and
XX phosphatases which are potentially useful as drug targets.
XX
PS Claim 1; Fig 5; Opp; English.
XX
XX The present invention relates to a method for detecting the presence of a
XX kinases or phosphatases encoded by nucleotides. The kinases and
XX phosphatases and their encoding nucleic acids are potentially useful as
XX drug targets. The present invention may also be useful in diagnosing
XX disease. The present sequence is human protein kinase CDNA
XX
SQ Sequence 361 BP; 71 A; 101 C; 120 G; 68 T; 0 U; 1 Other;
XX
Query Match 13.1%; Score 296.2; DB 10; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-46;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;
XX
QY 1613 CGAGATCTCGAAGCAAGTACAGCCACCCCAACATCGTGGCTCATTTGATGTCGACCC 1672
DB 20 CNAAGATCTCGAAGCAAGTACAGCCACCCCAACATCGTGGCTCATTTGATGTCGACCC 79
QY 1673 AGAAGCAGCCCATCTACATCGTCATGAGCTTGTGAGGGGGCGACTTCTGACCTTC 1732
DB 80 AGAAGCAGCCCATCTACATCGTCATGAGCTTGTGAGGGGGCGACTTCTGACCTTC 139
QY 1733 TCCGCACGAGGGGGCCCGCTCGGGGTGAAGACTCTGCTGCAGATGTGGGGATGCAG 1792
DB 140 TCCGCACGAGGGGGCCCGCTCGGGGTGAAGACTCTGCTGCAGATGTGGGGATGCAG 199
QY 1793 CTGCTGGCATGTGAGTCTCTGAGAGCAAGTCTGCATCCACCGGGACCTGGCTGCAG 1852
DB 200 CTGCTGGCATGTGAGTCTCTGAGAGCAAGTCTGCATCCACCGGGACCTGGCTGCAG 259
QY 1853 ACTGCTGTGACAGAGAAATGCTCTGAAG-ATCAGTACTTT-GGAGTGTCCGAGA 1910
DB 260 ACTGCTGTGACAGAGAAATGCTCTGAAGAAATCAGTACTTTGGGAGATGCCGAGA 319
QY 1911 GGAAGCC--GATGGGCTTATGCAAGCTTCAGGGGGGCTCAG 1949
DB 320 GGAAGCCGATTTGGGGTCTATGCAAGCTTCAGGGGGGCTCAG 361
XX
RESULT 15
ABX14976
ID ABX14976 standard; cDNA, 361 BP.
XX
AC ABX14976;
XX
DT 13-MAR-2003 (first entry)
```

```
XX
DE Human protein kinase CDNA mine16676human_sl.
XX
KW Human; ss; protein kinase; mine16676human_sl; cancer; cytostatic;
XX antianginal; hypotensive; cardiac; proliferative disorder;
XX cellular growth related disorder; cardiovascular disorder; heart failure;
XX hypertension; atrial fibrillation; dilated cardiomyopathy;
XX idiopathic cardiomyopathy; angina.
XX
OS Homo sapiens.
XX
PN US6465232-B1.
XX
PD 15-OCT-2002.
XX
PF 07-SEP-2001; 2001US-00948802.
XX
PR 31-AUG-1999; 99US-00387212.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Robison KE;
XX
DR WPI; 2003-147067/14.
XX
PT Novel human kinase and phosphatase nucleic acid molecules useful for
XX treating cellular proliferative disorders such as cancer, cardiovascular
XX diseases, hypertension, heart failure and angina.
XX
PS Claim 7; Fig 5; 47pp; English.
XX
XX The invention relates to an isolated human kinase and phosphatase nucleic
XX acid molecule appearing as ABX14972, ABX14973, ABX14975, ABX14976, and
XX ABX14978-ABX14981, or their complement. Also included are: (1) an
XX isolated nucleic acid molecule which is at least 90 % identical to the
XX nucleotide sequence ABX14973 or ABX14975, or 95 % identical to the
XX nucleotide sequence ABX14973, ABX14975, ABX14976, and ABX14978 or its
XX complement, where the nucleic acid molecule encodes a polypeptide having
XX a kinase activity; (2) An isolated nucleic acid molecule which hybridises
XX to ABX14973, ABX14975, ABX14976, and ABX14978 in 6X saline sodium citrate
XX (SSC) at 45 plusoc, followed by one or more washes in 0.2X SSC, 0.1%
XX sodium dodecyl sulphate (SDS) at 65 plusoc, where the molecule encodes a
XX polypeptide with kinase activity; (3) a vector comprising the
XX polynucleotides, operatively linked to a recombinant regulatory sequence;
XX and (4) expressing a polypeptide by culturing a host cell comprising the
XX vector under conditions in which the nucleic acid molecule is expressed.
XX The nucleic acids are useful as modulating agents in regulating a variety
XX of cellular processes, and fragments are useful as primers or
XX hybridisation probes for detecting kinase and phosphatase encoding
XX nucleic acids. The nucleic acids are useful for treating proliferative
XX disorders such as cancer and cellular growth related disorders including
XX cardiovascular disorders such as heart failure, hypertension, atrial
XX fibrillation, dilated cardiomyopathy, idiopathic cardiomyopathy or
XX angina. The present sequence is one of the human kinase or phosphatase
XX cDNAs of the invention
XX
SQ Sequence 361 BP; 71 A; 101 C; 120 G; 68 T; 0 U; 1 Other;
XX
Query Match 13.1%; Score 296.2; DB 10; Length 361;
Best Local Similarity 97.4%; Pred. No. 1.9e-46;
Matches 333; Conservative 0; Mismatches 4; Indels 5; Gaps 3;
XX
QY 1613 CGAGATCTCGAAGCAAGTACAGCCACCCCAACATCGTGGCTCATTTGATGTCGACCC 1672
DB 20 CNAAGATCTCGAAGCAAGTACAGCCACCCCAACATCGTGGCTCATTTGATGTCGACCC 79
QY 1673 AGAAGCAGCCCATCTACATCGTCATGAGCTTGTGAGGGGGCGACTTCTGACCTTC 1732
DB 80 AGAAGCAGCCCATCTACATCGTCATGAGCTTGTGAGGGGGCGACTTCTGACCTTC 139
QY 1733 TCCGCACGAGGGGGCCCGCTCGGGGTGAAGACTCTGCTGCAGATGTGGGGATGCAG 1792
DB 140 TCCGCACGAGGGGGCCCGCTCGGGGTGAAGACTCTGCTGCAGATGTGGGGATGCAG 199
```

```

QY 1793 CTGCTGCGATGAGTACCTGAGAGCAAGTGTGATCCACCGGGA CTTGGTGTCCGA 1852
DB 200 CTGCTGCGATGAGTACCTGAGAGCAAGTGTGATCCACCGGGA CTTGGTGTCCGA 259
QY 1853 ACTGCTGTGTGACAGAGAAATGTCTGAAAG-ATCAGTACTTT-GGATGTCCGAGA 1910
DB 260 ACTGCTGTGTGACAGAGAAATGTCTGAAAGATCAGTACTTTGGGGAATGTCCGAGA 319
QY 1911 GGAAGCC--GATGGGCTTATGCAAGCTCAGGGGGCTTCAG 1949
DB 320 GGAAGCCCGATTGGGGCTTATGCAAGCTCAGGGGGCTTCAG 361

```

Search completed: March 19, 2005, 04:31:26
 Job time : 1087.29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: March 19, 2005, 03:18:16 ; Search time 9383.68 Seconds
(without alignments)
11649.479 Million cell updates/sec

Title: US-10-660-763-1_COPY_72_2327

Perfect score: 2256

Sequence: 1 atggcgtctctctctgcgact.....gcattccgaagcgcgcgcgcg 2256

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: 1: gb_ba: 2: gb_ncg: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sbs: 12: gb_sy: 13: gb_un: 14: gb_vl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2256	100.0	2674	6	AR183261	AR183261 Sequence
2	2256	100.0	2674	6	AR456323	AR456323 Sequence
3	2036	90.2	2697	6	CQ725220	CQ725220 Sequence
4	2036	90.2	2788	9	BC035357	BC035357 Homo sapi
5	2029.6	90.0	2771	9	HSCFES	X52192 H. sapiens R
6	2023.4	89.7	2889	6	AX281847	AX281847 Sequence
7	1540.4	68.3	2680	10	WMPESCR	X12616 Mouse c-fes
8	1018.2	45.1	2715	14	ACFTS140A	M14930 Fujinami ba
9	1018.2	45.1	4788	14	ACR	J03194 Fujinami ba
10	1018.2	45.1	4788	14	AF033810	AF033810 Fujinami
11	841.4	37.3	4901	13	M22820	M22820 Figure 4. N
12	821.8	36.4	2397	14	FGSSTONC	J02088 Feline sarc
13	810.8	35.9	2946	14	FGGGAONC	J02087 feline sarc
14	656.4	29.1	4788	5	BC073445	BC073445 Xenopus l
15	635	28.1	856	6	A06935	A06935 feline sarc
16	626.2	27.8	2284	14	ACGAGFPs	K01690 Avian sarc
17	443	19.6	772	10	BC038130	BC038130 Mus muscu
18	376.6	16.7	1563	10	BC051249	BC051249 Mus muscu
19	376.6	16.7	2948	10	BC058100	BC058100 Mus muscu

20	376.6	16.7	2994	10	MMU76762	U76762 Mus musculu
21	375	16.6	2069	10	MUSP8RT	M32054 Mouse tyros
22	374.8	16.6	1171	10	RNPLK	X13412 Rat mRNA fo
23	342.8	15.2	2947	4	AF187884	AF187884 Canis fam
24	328.8	14.6	2950	9	HUMTKFER	J03358 Human tyros
25	318.4	14.1	3882	3	SREPSTR	Y17051 Sycon rapha
26	296.2	13.1	361	6	AR175607	AR175607 Sequence
27	296.2	13.1	361	6	AR236724	AR236724 Sequence
28	274.4	12.2	2786	3	ABO06567	ABO06567 Ephratia
29	271.6	12.0	3198	6	CQ597087	CQ597087 Sequence
30	271.6	12.0	3369	6	CQ591168	CQ591168 Sequence
31	271.6	11.9	4972	3	BT003462	BT003462 Drosophila
32	268.4	12.0	2303	3	DMU50450	U50450 Drosophila
33	266.8	11.8	3297	3	DMPS85D	X52844 Drosophila
34	230.6	10.2	2597	6	CQ726604	CQ726604 Sequence
35	227.2	10.1	15297	6	AR183262	AR183262 Sequence
36	227.2	10.1	15297	6	AR456324	AR456324 Sequence
37	225.6	10.0	12263	9	HSPESFPS	X06592 Human c-fes
38	224.6	10.0	142301	2	AC124248	AC124248 Homo sapi
39	224.6	10.0	156889	2	AC004586	AC004586 Homo sapi
40	211.4	9.4	4160	6	A06936	A06936 H. sapiens E
41	211.4	9.4	4160	6	E01406	E01406 DNA sequenc
42	201.2	8.9	2398	9	HSRTKEPH	Z27409 H. sapiens m
43	198	8.8	2955	6	AX481387	AX481387 Sequence
44	198	8.8	3370	6	CQ776660	CQ776660 Sequence
45	198	8.8	3370	6	ARS59315	ARS59315 Sequence

ALIGNMENTS

RESULT 1	AR183261	2674 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR183261	Sequence 1 from patent US 6340584.			
DEFINITION	Sequence 1 from patent US 6340584.				
ACCESSION	AR183261				
VERSION	AR183261.1	GI:20226854			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 2674)				
AUTHORS	Gan, W., Ye, J., Di Francesco, V. and Beasley, E. M.				
TITLE	Isolated human kinase proteins, and uses thereof				
JOURNAL	Patent: US 6340584-A 1 22-JAN-2002;				
FEATURES	Location/Qualifiers				
source	1..2674				
	/organism="unknown"				
	/mol_type="unassigned DNA"				
ORIGIN					
Query Match	100.0%; Score 2256; DB 6; Length 2674;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 2256; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	ATGGGCTTCTCTTCTGAGCTGTGACGCCCGGAGGCGACGGGGTCTCTGACGAAATGCG 60			
DB	72	ATGGGCTTCTCTTCTGAGCTGTGACGCCCGGAGGCGACGGGGTCTCTGACGAAATGCG 131			
QY	61	GAGGCGGACCTTCCTTACTGAGGCGGACGAAAGTGATGGCCGAGGGGTCAAGAGT 120			
DB	132	GAGGCGGACCTTCCTTACTGAGGCGGACGAAAGTGATGGCCGAGGGGTCAAGAGT 191			
QY	121	GACAGGAGTATGACGAGCTGCTTCAACATGTCCTGACGAGCAGTGGGGCCAGAGC 180			
DB	192	GACAGGAGTATGACGAGCTGCTTCAACATGTCCTGACGAGCAGTGGGGCCAGAGC 251			
QY	181	CGAGGCATGACGCGCTTGACAGCCCATCATGCTGCTGGGCTGAGATCAACAGCCAACT 240			
DB	252	CGAGGCATGACGCGCTTGACAGCCCATCATGCTGCTGGGCTGAGATCAACAGCCAACT 311			
QY	241	GAGGCGCTGAGCGGCTTGCTGGGCGACGACGAGAGATCTGAACCTCAGGGCCCTGAGC 300			

312 GAGGGCTGAGCGCTTGGCTGGGGGAGCAAGAGAGATCTGAATCAAGGGCCCTCGAGC 371
301 AAGCTGAGCTGCTGATTCGGGAAACGGCAGAGCTTTCGAAAGCTTACAGCGAGCAGTGG 360
372 AAGCTGAGCTGCTGATTCGGGAAACGGCAGAGCTTTCGAAAGCTTACAGCGAGCAGTGG 431
361 CAGCAGCTGAGAGAGAGCTCAACCAAGACCCAGCAGGAGCAATTGAGAGCTGAAGAGC 420
432 CAGCAGCTGAGAGAGAGCTCAACCAAGACCCAGCAGGAGCAATTGAGAGAGCTGAAGAGC 491
421 CAGTACCGAGCTTGGCAAGGGAACAGTGGCCCAAGCCAGGAGCAAGTACCAAGAGGCGAC 480
492 CAGTACCGAGCTTGGCAAGGGAACAGTGGCCCAAGCCAGGAGCAAGTACCAAGAGGCGAC 551
481 AAGAGCAAGAGCCGTGACAAAGGCAAGCAAGTATGTGGCAGCCTTGGAAAGCTCTTT 540
552 AAGAGCAAGAGCCGTGACAAAGGCAAGCAAGTATGTGGCAGCCTTGGAAAGCTCTTT 611
541 GCTCACCAACACGCTATGTGCTGGGGGAGCGGGCTGGCAGCTACACCAAGCAGCAGC 600
612 GCTCACCAACACGCTATGTGCTGGGGGAGCGGGCTGGCAGCTACACCAAGCAGCAGC 671
601 CACCAAGCTCTGCTGCGCCGCGCTGCTGCGGTCACTGCAAGACCTGCAAGAGAGATGGCT 660
672 CACCAAGCTCTGCTGCGCCGCGCTGCTGCGGTCACTGCAAGACCTGCAAGAGAGATGGCT 731
661 TGGATCTTGAAGAGATTCCTGCAAGAAATACCTGGAATTTAGCAGCCTTGGTGGAGATGAG 720
732 TGGATCTTGAAGAGATTCCTGCAAGAAATACCTGGAATTTAGCAGCCTTGGTGGAGATGAG 791
721 GTGGTGGCCATTTACCGGGAGATGGCTGAGCTGCTGCGCCGCTGCAAGCCTGAGGCTGAG 780
792 GTGGTGGCCATTTACCGGGAGATGGCTGAGCTGCTGCGCCGCTGCAAGCCTGAGGCTGAG 851
781 TACCAAGGCTTCTGCGCAAGTATGGGTCCGCACTGACGTCCACCTGTGTCAAGCTTC 840
852 TACCAAGGCTTCTGCGCAAGTATGGGTCCGCACTGACGTCCACCTGTGTCAAGCTTC 911
841 GATGATCTGCTGTTGAGAGAGGGTGAACCGCTGGAACCTTGGGAGAGCTTCCAGCTGAACGAG 900
912 GATGATCTGCTGTTGAGAGAGGGTGAACCGCTGGAACCTTGGGAGAGCTTCCAGCTGAACGAG 971
901 CTGACCTGAGAGCGCTGAGCAACGAGCTGACCTGAGTGAAGATGAGCTGCTGAGCC 960
972 CTGACCTGAGAGCGCTGAGCAACGAGCTGACCTGAGTGAAGATGAGCTGCTGAGCC 1031
961 ACCGAGATGCTGTTCAAGCGGAGAGAGATGGTTACGACGCTGCAACAGAGCTCCGGAAT 1020
1032 ACCGAGATGCTGTTCAAGCGGAGAGAGATGGTTACGACGCTGCAACAGAGCTCCGGAAT 1091
1021 GAAGAAGAGAAACCCACCCCCCGGAGCGGGTGAAGCTGCTGGGCAAGAGCAAGTGTCTG 1080
1092 GAAGAAGAGAAACCCACCCCCCGGAGCGGGTGAAGCTGCTGGGCAAGAGCAAGTGTCTG 1151
1081 CAAGAAGCACTGAGAGGGGCTGCAAGTATGAGCTGAGCAAGGCAAGAGCTGAGCCAG 1140
1152 CAAGAAGCACTGAGAGGGGCTGCAAGTATGAGCTGAGCAAGGCAAGAGCTGAGCCAG 1211
1141 CAGAGATGCTGAGACCAAGGCTGAGCACTGAGGCGCCGAGCGAGCCCTGTGTCTG 1200
1212 CAGAGATGCTGAGACCAAGGCTGAGCACTGAGGCGCCGAGCGAGCCCTGTGTCTG 1271
1201 CTCTCTGAGATACCGCCCACTCCAGTGTCTCTCGGAGAGAGAGAGAGAGGGGGGAGAG 1260
1272 CTCTCTGAGATACCGCCCACTCCAGTGTCTCTCGGAGAGAGAGAGAGGGGGGAGAG 1331
1261 ACACCAAGCTGAGAGATCTTAAAGAGCAATCTCAAGAAATCTTCCGCCCAAGTGTCTG 1320
1332 ACACCAAGCTGAGAGATCTTAAAGAGCAATCTCAAGAAATCTTCCGCCCAAGTGTCTG 1391
1321 AACCTGACCACTGGAAGGGGAGAGGCTTTCAGCAATTCCTTGTCTCATGACCACTTA 1380

1392 AACCTGACCACTGGAAGGGGAGAGGCTTTCAGCAATTCCTTGTCTCATGACCACTTA 1451
1381 CTGAGACACCCAGACAGCCCTCTCAACAAAGAGTGGTGTCTCTGACAGAGGCTGTGCC 1440
1452 CTGAGACACCCAGACAGCCCTCTCAACAAAGAGTGGTGTCTCTGACAGAGGCTGTGCC 1511
1441 AAGAGCAAGTGGGTCTGAAACATGAGAGCTGTGTGGGTGAGCAGATTTGAGCGGGG 1500
1512 AAGAGCAAGTGGGTCTGAAACATGAGAGCTGTGTGGGTGAGCAGATTTGAGCGGGG 1571
1501 AACTTGGGGAAGTGTTCAGCGGAGCGCTTGGCAAGCCCAACACCTTGGTGGCGTGAAG 1560
1572 AACTTGGGGAAGTGTTCAGCGGAGCGCTTGGCAAGCCCAACACCTTGGTGGCGTGAAG 1631
1561 TCTTGTGAGAGAGCGCTCCACCTGACCTCAAGGCAAGTTTCTACAGAGAGAGAGATC 1620
1632 TCTTGTGAGAGAGCGCTCCACCTGACCTCAAGGCAAGTTTCTACAGAGAGAGAGATC 1691
1621 CTGAAGCAGTACAGCCACCCCAACATCGTCTCTCATTTGTGTGCAACCAAGAGCAG 1680
1692 CTGAAGCAGTACAGCCACCCCAACATCGTCTCTCATTTGTGTGCAACCAAGAGCAG 1751
1681 CCCATCTACATGTGATGAGAGCTTGTGCAAGGGGGGCGAATTTCTGACCTTCCGCAAG 1740
1752 CCCATCTACATGTGATGAGAGCTTGTGCAAGGGGGGCGAATTTCTGACCTTCCGCAAG 1811
1741 GAGGGGGGCGCGCTGCGGGTGAAGACTGTGCTGCAAGATGTTGGGGGATGACAGCTGTGCG 1800
1812 GAGGGGGGCGCGCTGCGGGTGAAGACTGTGCTGCAAGATGTTGGGGGATGACAGCTGTGCG 1871
1801 ATGAGATACCTGAGAGAGCAAGTGTGATCAACCGGGAACCTTGGCTGAGAACTGCTTG 1860
1872 ATGAGATACCTGAGAGAGCAAGTGTGATCAACCGGGAACCTTGGCTGAGAACTGCTTG 1931
1861 GTGACAGAGAGAAATGTCTGAAAGATCAAGTGAATTTGGGATGTTCCGAGAGAGAGCCGAT 1920
1932 GTGACAGAGAGAAATGTCTGAAAGATCAAGTGAATTTGGGATGTTCCGAGAGAGAGCCGAT 1991
1921 GGGGTCTATGCAAGCGCTCAAGGGGGGCTCAGACAAGTCCCCTGTAAGTGAACCGCACTTGAG 1980
1992 GGGGTCTATGCAAGCGCTCAAGGGGGGCTCAGACAAGTCCCCTGTAAGTGAACCGCACTTGAG 2051
2052 GCCCTTAACTACAGGCGGCTACTCTCCGAAAGCAGAGTGTGAGCTTGGCATCTTGCTC 2111
2041 TGGAGAGCTTACAGCTTGGGGGCTTCCCTATATCCCACTGAGCAATGAGCAGACAGG 2100
2112 TGGAGAGCTTACAGCTTGGGGGCTTCCCTATATCCCACTGAGCAATGAGCAGACAGG 2171
2101 GAGTTTGTGAGAGAGGGGGGCGGTCTGACCTGACCCAGAGCTGTGTCTGATGCGGTGTC 2160
2172 GAGTTTGTGAGAGAGGGGGGCGGTCTGACCTGACCCAGAGCTGTGTCTGATGCGGTGTC 2231
2161 AGGCTCATAGAGCAGTGTGAGGCTATGAGCCTTGGGAGCGGCGCCAGCTTCAAGCAATC 2220
2232 AGGCTCATAGAGCAGTGTGAGGCTATGAGCCTTGGGAGCGGCGCCAGCTTCAAGCAATC 2291
2221 TACCAAGAGCTGAGAGCATCCGAAAGCGGCAATCGG 2256
2292 TACCAAGAGCTGAGAGCATCCGAAAGCGGCAATCGG 2327

RESULT 2
AR456323 2674 bp DNA linear PAT 20-FEB-2004
LOCUS AR456323
DEFINITION Sequence 1 from patent US 6686187.
ACCESSION AR456323
VERSION AR456323.1 GI:42691378
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 2674)
AUTHORS Gan, W., Ye, J., Di Francesco, V. and Beasley, E. M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding
JOURNAL Patent: US 6686187-A 1 03-FEB-2004;
FEATURES Location/Qualifiers
source 1. 2674
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 2256; DB 6; Length 2674;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 60
2 ATGGGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 131
61 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 120
132 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 191
121 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 180
192 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 251
181 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 240
252 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 311
241 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 300
312 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 371
301 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 360
372 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 431
361 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 420
432 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 491
421 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 480
492 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 551
481 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 540
552 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 611
541 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 600
612 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 671
601 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 660
672 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 731
661 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 720
732 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 791
721 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 780
792 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 851
781 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 840
852 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 911
841 GAGGCGGAGCTTCTCTCTGAGCTGTGCAAGCCCAAGGCAAGGAGTCTCTGCAAGCAATGACG 900

912 GATGAGTCACTGCTTGAAGAGGAGTGAACCGCTGAGCCTGAGGAGCTCCAGCTGAACGAG 971
901 CTGAGCTGTGAGAGGCTGTGAGCAACAGCTGACCTTCAAGTGAAGTGAAGTGTGAGCC 960
972 CTGAGCTGTGAGAGGCTGTGAGCAACAGCTGACCTTCAAGTGAAGTGAAGTGTGAGCC 1031
961 ACCGAGATGATGTTCAAGGCGGAGGAGGATGTTTACGAGCTGCAACAGGAGCTCCGGAAT 1020
1032 ACCGAGATGATGTTCAAGGCGGAGGAGGATGTTTACGAGCTGCAACAGGAGCTCCGGAAT 1091
1021 GAAAGAGAGAACACCCACCCCGGAGCGGAGTGAAGCTGCTGAGGCAAGAGCAAGTGTG 1080
1092 GAAAGAGAGAACACCCACCCCGGAGCGGAGTGAAGCTGCTGAGGCAAGAGCAAGTGTG 1151
1081 CAAGAGCACTGAGGAGGCTGCAAGTGAAGCTGTGCAAGGCAAGGCTGTGCAAGGCTGAG 1140
1152 CAAGAGCACTGAGGAGGCTGCAAGTGAAGCTGTGCAAGGCAAGGCTGTGCAAGGCTGAG 1211
1141 CAAGAGTGTGCAAGGAGGCTGCAAGTGAAGCTGTGCAAGGCAAGGCTGTGCAAGGCTGAG 1200
1212 CAAGAGTGTGCAAGGAGGCTGCAAGTGAAGCTGTGCAAGGCAAGGCTGTGCAAGGCTGAG 1271
1201 CTCTGTGAGAGTGAAGGAGGCTTCAAGTGTGCTGCTGAGCAAGAGGAGGAGGAGGAG 1260
1272 CTCTGTGAGAGTGAAGGAGGCTTCAAGTGTGCTGCTGAGCAAGAGGAGGAGGAGGAGG 1331
1261 ACAACCAAGCTGAGAGTCTTGAAGAGCAATCTCAAGATCTTCCGCCCCAAGTCTGAG 1320
1332 ACAACCAAGCTGAGAGTCTTGAAGAGCAATCTCAAGATCTTCCGCCCCAAGTCTGAG 1391
1321 AACCTGTACCACTGAGAGGAGGAGGCTTCCAGATCTCTTGTGCTCAATGAGCAACCTA 1380
1392 AACCTGTACCACTGAGAGGAGGAGGCTTCCAGATCTCTTGTGCTCAATGAGCAACCTA 1451
1381 CTGAGCAACCAAGGAGGAGGCTTCCAGATCTCTTGTGCTCAATGAGCAACCTA 1440
1452 CTGAGCAACCAAGGAGGAGGCTTCCAGATCTCTTGTGCTCAATGAGCAACCTA 1511
1441 AAGAGCAAGTGAAGTGTGAACATGAGAGCTGTGTGTGAGAGCAATGAGAGGAGGAG 1500
1512 AAGAGCAAGTGAAGTGTGAACATGAGAGCTGTGTGTGAGAGCAATGAGAGGAGGAG 1571
1501 AACTTTGGGAGAGTGTGAACATGAGAGCTGTGTGTGAGAGCAATGAGAGGAGGAG 1560
1572 AACTTTGGGAGAGTGTGAACATGAGAGCTGTGTGTGAGAGCAATGAGAGGAGGAG 1631
1561 TCTTGTGAGAGAGCTTCCCACTGAGCTCAAGGCAAGTGTGTGAAGAGGAGGAGT 1620
1632 TCTTGTGAGAGAGCTTCCCACTGAGCTCAAGGCAAGTGTGTGAAGAGGAGGAGT 1691
1621 CTGAGCAAGTGAAGGAGGAGGCTTCCCACTGAGCTCAAGGCAAGTGTGTGAAGAGGAG 1680
1692 CTGAGCAAGTGAAGGAGGAGGCTTCCCACTGAGCTCAAGGCAAGTGTGTGAAGAGGAG 1751
1681 CCATCTAATGATGTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
1752 CCATCTAATGATGTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1811
1741 GAG 1800
1812 GAG 1871
1801 ATGAGTACCTGAGAGAGCAAGTGTGATTCACCGGAGCTTGAAGCTTCCGGAATCTGCTG 1860
1872 ATGAGTACCTGAGAGAGCAAGTGTGATTCACCGGAGCTTGAAGCTTCCGGAATCTGCTG 1931
1861 GTGAGCAAGAGAGAGTGTCTTGAAGAGCAAGTGTGATTCACCGGAGCTTGAAGCTTCCG 1920
1932 GTGAGCAAGAGAGAGTGTCTTGAAGAGCAAGTGTGATTCACCGGAGCTTGAAGCTTCCG 1991
1921 GAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
1992 GAGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2051

Qy	1981	GCCCTTAACATAGGCGCGGTACTCCTCCGCAGAAAGCAAGCATGGAGCTTTGGGATCTGTC	2048
Db	2052	GCCCTTAACTACGGCCGCTACTCTCTCCGAAGCAAGTGTAAGCTTTGGCATCTTGCTC	2111
Qy	2041	TGGAGAACCTTTCAGGCTTGSGGGGCTCCGCCCTATCCCAACTTCAGCAATCAGCAGAACAAG	2100
Db	2112	TGGAGAGACCTTCAGGCTTGSGGGGCTCCGCCCTATCCCAACTTCAGCAATCAGCAGAACAAG	2171
Qy	2101	GAGTTTGTGAGAAAGGGGGGCGCTCTGCGCTGCCACAGAGCTGTGTCTGATGCGCGTGTTC	2160
Db	2172	GAGTTTGTGAGAAAGGGGGGCGCTGTGCGCTGCCACAGAGCTGTGTCTGATGAGCGGTGTTC	2231
Qy	2161	AGGCTCATGAGACAGTGTCTGGGCGCTATGAGCGCTGGGCAAGCGCCAGGCTTCAGACACATC	2220
Db	2232	AGGCTCATGAGACAGTGTCTGGGCGCTATGAGCGCTGGGCAAGCGCCAGGCTTCAGACACATC	2291
Qy	2221	TACCAGAGAGCTGCAGACGATCCGAAAGGGGCGATCGG	2256
Db	2292	TACCAGAGAGCTGCAGAGCATCCGAAAGGGGCGATCGG	2327

RESULT 3				
CQ725220				
LOCUS	CQ725220	2697 bp	DNA	linear
DEFINITION	Sequence 11154 from Patent WO02068579.			
ACCESSION	CQ725220			
VERSION	CQ725220.1	GI:42286077		
KEYWORDS				
SOURCE	Homo sapiens	(human)		
ORGANISM	Homo sapiens			

REFERENCE	AUTHORS	TITLE
1	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.	Kits, such as nucleic acid arrays, comprising a majority of

JOURNAL
Patent: WO 02068579-A 11154 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES	Location/Qualifiers
source	1..2697
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"

ORIGIN

Query Match	90.2%	Score 2036;	DB 6;	Length 2697;
Best Local Similarly	91.5%;	Pred. No. 0;		
Matches 2356; Conservative	0;	Mismatches	210;	Gaps 1;

Oy	1	ATGGGCTCTCTTCTGAGCTGTGACGCCGCCAGAGGCCACAGGGAGTCTGACGAATGAG	60
Db	11	ATGGGCTTCTCTTCTGAGCTGTGACGCCGCCAGAGGCCACAGGGAGTCTGACGAATGAG	70
Oy	61	GAGGCCGAGACTTCTGTCTACTGAGAGGCGATGAGAAATGATATGCCCCAGGGATCAAGAT	120
Db	71	GAGGCCGAGCTTCTGTCTACTGAGAGGCGATGAGAAATGATATGCCCCAGGGATCAAGAT	130
Oy	121	GACAGGGAGTATGACGAGACTGCTTACCAACATGTCTCTGACAGGACATGTGGGGCCAGAGC	180
Db	131	GACAGGGAGTATGACGAGACTGCTTACCAACATGTCTCTGACAGGACATGTGGGGCCAGAGC	190
Oy	181	CGGGGCATCAGCCCTTGACAGACCCCATTCAGTCAGTCTTGGGCTGAGATATCAACAGCCAACT	240
Db	191	CGGGGCATCAGCCCTTGACAGACCCCATTCAGTCAGTCTTGGGCTGAGATATCAACAGCCAACT	250
Oy	241	GAGGGCCTGAGCCGCTTGTCTGTGGGAGAGACGACAGAGATTCGAATTCAGGGCCCCCTGAGC	300
Db	251	GAGGGCCTGAGCCGCTTGTCTGTGGGAGAGACGACAGAGATTCGAATTCAGGGCCCCCTGAGC	310
Oy	301	AAGCTGAGCCTGTCTCATCCGGGAAACGGGACAGCACTTTGCAAGCACTTACAGGAGACATGTG	360

Db	311	AAAGTGAGCCTGCTCATCCGGGAAAGGCAAGAGTTTGCAAAACCTTACAGGAGCAGTTG	370
OY	361	CAGCAGCTGCACAGAGGCTTCAACAGACCCAACGCCAGGACATTGGAAGCTGAGAGC	420
Db	371	CAGCAGCTGCACAGAGGCTTCAACAGACCCAACGCCAGGACATTGGAAGCTGAGAGC	430
OY	421	CAGAACGAGGTCCTGGACGGGAGACAGTGGCCAAAGCCAAAGGCAAGTACAGAGAGCCAGC	480
Db	431	CAGTACCGAGCTCTGGACCGGAGACAGTGGCCAAAGGCAAGTACAGAGAGCCAGC	490
OY	481	AAAGCAAGGACCGTGCACAAAGGCCAAGAGACAAATATGTGCGCAGCCTGTGAACTCTTT	540
Db	491	AAAGCAAGGACCGTGCACAAAGGCCAAGAGACAAATATGTGCGCAGCCTGTGAACTCTTT	550
OY	541	GCTACCAACAAACCGCTATGTGTCTGGGGGTGCGGGCTGGCGAGCTAACACACACAGCACAC	600
Db	551	GCTACCAACAAACCGCTATGTGTCTGGGGGTGCGGGCTGGCGAGCTAACACACAGCACAC	610
OY	601	CACCAAGCTCTGCTGCGCCCGGCTGCTCGGTCTCATCTGAGAGACCTGCACAGAGAGATGGT	660
Db	611	CACCAAGCTCTGCTGCGCCCGGCTGCTCGGTCTCATCTGAGAGACCTGCACAGAGAGATGGT	670
OY	661	TGCATCTCTGAAAGAGATCTTCAGGAATTACCTTGAGATTATGCAAGCTTGTGTGAGAGATGAG	720
Db	671	TGCATCTCTGAAAGAGATCTTCAGGAATTACCTTGAGATTATGCAAGCTTGTGTGAGAGATGAG	730
OY	721	GTGTGTGCGCAATTCACCGGAGAGATGGCTGTGCAGCTGCTCCGCAATCCAGCCTGAGAGCTGAG	780
Db	731	GTGTGTGCGCAATTCACCGGAGAGATGGCTGTGCAGCTGCTCCGCAATCCAGCCTGAGAGCTGAG	790
OY	781	TACCAAGGCTTCTGTGCGACAGTATGGTTCGCGACTTGACGTCCTCCACTGTGTTCACGTTT	840
Db	791	TACCAAGGCTTCTGTGCGACAGTATGGTTCGCGACTTGACGTCCTCCACTGTGTTCACGTTT	850
OY	841	GATGATGCTCATGCTGTGAGAGAGGGTGAACCGCTGGAGCTGGGGAGGCTTCAAGCTGAAGAG	900
Db	851	GATGATGCTCATGCTGTGAGAGAGGGTGAACCGCTGGAGCTGGGGAGGCTTCAAGCTGAAGAG	910
OY	901	CTGACTGTGGAGAGCGTGCAGACACCGCTGACCTCACTGACAGATGACAGATGAGCTTGTTGGCC	960
Db	911	CTGACTGTGGAGAGCGTGCAGACACCGCTGACCTCACTGACAGATGAGCTTGTTGGCC	970
OY	961	ACCGAGATGTGTTTCAAGCGCGCAGAGAAATGTTTACGACGTGCACACAGAGCTTCCGAAT	1020
Db	971	ACCGAGATGTGTTTCAAGCGCGCAGAGAAATGTTTACGACGTGCACACAGAGCTTCCGAAT	1030
OY	1021	GAAGAGGAGAAACACCAACCCCGGGAGCGGGGTGACGTGCTGGGCAAGAGCAAGTCTGTG	1080
Db	1031	GAAGAGGAGAAACACCAACCCCGGGAGCGGGGTGACGTGCTGGGCAAGAGCAAGTCTGTG	1090
OY	1081	CAAGAGCACTGCAGGGGCTGCAAGTATGCGTGTGCACGCCAGCCAAAGCTTGCAGGCCAG	1140
Db	1091	CAAGAGCACTGCAGGGGCTGCAAGTATGCGTGTGCACGCCAGCCAAAGCTTGCAGGCCAG	1150
OY	1141	CAGAGATTGCTGCAGACCAAGCTGGGACACCTGGGGCCCGGGCAGCCCCCGCTGTGCTG	1200
Db	1151	CAGAGATTGCTGCAGACCAAGCTGGGACACCTGGGGCCCGGGCAGCCCCCGCTGTGCTG	1210
OY	1201	CTTCTGCAGATGACCGCACTCCACGTCGTCTCTCGAGCAGAGCGAGAGGGGGAGAG	1260
Db	1211	CTTCTGCAGATGACCGCACTCCACGTCGTCTCTCGAGCAGAGCGAGAGGGGGAGAG	1270
OY	1261	ACACCCACGCTGGAGATCTTTAAGAGCCACATCTCAGGAATCTTCCCGCCCCCAAGTTCTCG	1320
Db	1271	ACACCCACGCTGGAGATCTTTAAGAGCCACATCTCAGGAATCTTCCCGCCCCCAAGTTCTCG	1330
OY	1321	-----	1320
Db	1331	CTCCCTCAACGCTGCAGCTCATTTCCGGAGGTGCAAGACCCCTGCATGAGCAAGCTGTGG	1390
OY	1321	-----	1320
Db	1391	TACCAACGGGCGCATCCCGAGGGCAGAGGTGGCTGAGCTGCTGTGCACTCTGGGGAATTTC	1450

gene
1. .2788
/gene="FES"
/note="synonym: FES"
/db_xref="LocustID:2242"
/db_xref="MTM:190030"
73. .2541
/gene="FES"
/codon_start=1
/product="V-FES feline sarcoma viral/V-FPS fujinami avian
sarcoma viral oncogene homolog"
/protein_id="AAH3537.1"
/db_xref="GI:23271525"
/db_xref="LocustID:2242"
/db_xref="MTM:190030"
/translation="MGFSEELGSGPOHGVILQMOEAEILRLSGMKMAORYKPREY
AULHHMSLDGSGGSRATSPSPISQSAEITISQTEGSLRLROAEDLNGPSKL
SLIIRKQRLRTYSEWQLOOELTKTISODI EKLSQYRLARSAQAKRYDAS
KQKDRKADKRYRSLMKLFHHNRVLGVRAQLHQRHQLLFGRLSDYDHEE
MACILKEILQBYLEISLVQDEVAIHRMAAARIQDEAYOGLRQYGSAPVPP
CTPDESLEEGSPLEPGLQINELTVEVOHTLTSTDELAVALATMVPFRDMVTL
QOELRREENTHREHREVOILGKROYALQGLQVALGSAQALQOELLQRLHEIG
RBPRLVLLQDRHSTSSSEDRSGRPTLEILKSHISGI PRPPESL.PPRLQILPE
VQPLHEQLMTHGATPRAEVAELVHSGPLVRESQKQBYLSVLMDELPHFTIIS
LNLVRLGEGSPSIPLLIDHLSTQPLTKSGVLAHRAVPDXLVNHEDLVAGEQ
IGNGEVEFSGRLRADNTLVAVKSCRETL.PDLKAKFLQEARILKQYSPNIVRLIG
VCTQKQPIYIVNELVQGDPLFLRTGEGARLVKTLQVGDAAAGAELESKCIHR
DLAARCLVTEKRVKISDFGMSREEDGVAAAGSLRQVPYKTPAPEALANTGRYSSE
SPVMSRGLIMETPSLGAAPYPLSMQOTREFPEKGRPLPCELCPDAVFLMEQDMA
YEPGORPSFTTYQELQSTIRKRR"

ORIGIN

Query Match 90.2%; Score 2036; DB 9; Length 2788;

Best Local Similarity 91.5%; Pred. No. 0;

Matches 2256; Conservative 0; Mismatches 0; Indels 210; Gaps 1;

QY 1 ATGGGCTTCTCTTCTGAGCTGTGACAGCCCGACGGGCGACGGGCTCTTGCAGCAATGCAAG 60
DB 73 ATGGGCTTCTCTTCTGAGCTGTGACAGCCCGACGGGCGACGGGCTCTTGCAGCAATGCAAG 132
QY 61 GAGGCCGAGCTTGTCTAAGGAGGCGATGAGAAAGTGAATGGCCGACGGGCTCAAGAGT 120
DB 133 GAGGCCGAGCTTGTCTAAGGAGGCGATGAGAAAGTGAATGGCCGACGGGCTCAAGAGT 192
QY 121 GACAGGAGATGATCAGAGACTGCTTCAACAATGTCCCTGACAGGACAGTGGGGCCAGAGC 180
DB 193 GACAGGAGATGATCAGAGACTGCTTCAACAATGTCCCTGACAGGACAGTGGGGCCAGAGC 252
QY 181 CGGGCCATCAGCCCTGACAGCCCATCACTCACTCTGGGCTGAGATCACAGCCAAACT 240
DB 253 CGGGCCATCAGCCCTGACAGCCCATCACTCACTCTGGGCTGAGATCACAGCCAAACT 312
QY 241 GAGGGCTGAGCCGCTTGTCTGCGGCGACAGGACGAGAGATGTAATCAAGGGCCCTGAGC 300
DB 313 GAGGGCTGAGCCGCTTGTCTGCGGCGACAGGACGAGAGATGTAATCAAGGGCCCTGAGC 372
QY 301 AAGCTGAGCTGCTCAATCCGGGAAACGGACAGACTTGCAGAGACTTACAGGACAGTGG 360
DB 373 AAGCTGAGCTGCTCAATCCGGGAAACGGACAGACTTGCAGAGACTTACAGGACAGTGG 432
QY 361 CAGCAGCTGACAGAGAGCTACCAAGACCACAGCCAGGACATTGAGAAAGCTGAAAGC 420
DB 433 CAGCAGCTGACAGAGAGCTACCAAGACCACAGCCAGGACATTGAGAAAGCTGAAAGC 492
QY 421 CAGTACGAGCTGAGAGGGGACAGTGGCCAGGACCAAGGCAAGTACCAAGAGGGCCAGC 480
DB 493 CAGTACGAGCTGAGAGGGGACAGTGGCCAGGACCAAGGCAAGTACCAAGAGGGCCAGC 552
QY 481 AAAGACAAGGACCGTGAACAAGCCCAAGGACAAATATGTGGCAGCCCTGTGAAAGCTTTT 540
DB 553 AAAGACAAGGACCGTGAACAAGCCCAAGGACAAATATGTGGCAGCCCTGTGAAAGCTTTT 612
QY 541 GCTTACCAACAACGCTATGTGTGGGCTGGGGCTGCGGAGTTACCAACAACAACAC 600

DB 613 GCTCAACAACAACGCTATGTGTGGGCTGGGGGCTGCGACACTACCAACCAAGACACAC 672
QY 601 CACAGAGCTCTGCTGCGCGGCTGCTGCGGCTCACTGACAGGACCTGCAACGAGAGATGCT 660
DB 673 CACAGAGCTCTGCTGCGCGGCTGCTGCGGCTCACTGACAGGACCTGCAACGAGAGATGCT 732
QY 661 TGCATCTGAAAGAGATCTGCGAGGAATACCTGAGATTAGAGCTGTGTGACAGATGAG 720
DB 733 TGCATCTGAAAGAGATCTGCGAGGAATACCTGAGATTAGAGCTGTGTGACAGATGAG 792
QY 721 GTGTGTGCGCATTCACCGGAGATGTGCTGAGAGTGTGCTCCCGGATCCAGGCTGAGTGA 780
DB 793 GTGTGTGCGCATTCACCGGAGATGTGCTGAGAGTGTGCTCCCGGATCCAGGCTGAGTGA 852
QY 781 TACCAAGGCTTCTGCGACAGATGAGTGGTCCGACCTGACGTCCACCTGTGTACAGTTT 840
DB 853 TACCAAGGCTTCTGCGACAGATGAGTGGTCCGACCTGACGTCCACCTGTGTACAGTTT 912
QY 841 GATGAGTCACTGCTTGAAGAGGGTGAACCGCTGAGAGCTGGGGAGCTCCAGCTGAACGAG 900
DB 913 GATGAGTCACTGCTTGAAGAGGGTGAACCGCTGAGAGCTGGGGAGCTCCAGCTGAACGAG 972
QY 901 CTGACTGTGAGAGAGGTGACAGCAACGTGACCTGACATGACAGATGAGTGTGTGGGCC 960
DB 973 CTGACTGTGAGAGAGGTGACAGCAACGTGACCTGACATGACAGATGAGTGTGTGGGCC 1032
QY 961 ACCGAGATGATGTTGACGGCGGACGAGAGATGTTACGCACTGCAACAGGAGCTCCGGAAT 1020
DB 1033 ACCGAGATGATGTTGACGGCGGACGAGAGATGTTACGCACTGCAACAGGAGCTCCGGAAT 1092
QY 1021 GAAAGAGAGAACCAACCAACCCCGGAGACGGGTGACAGCTGTGGGCAAGAGCAATGCTTG 1080
DB 1093 GAAAGAGAGAACCAACCAACCCCGGAGACGGGTGACAGCTGTGGGCAAGAGCAATGCTTG 1152
QY 1081 CAAGAAGCACTGACAGGGCTGAGAGTGAAGGCTGTGACAGCCAGAGCTGACAGGCCAG 1140
DB 1153 CAAGAAGCACTGACAGGGCTGAGAGTGAAGGCTGTGACAGCCAGAGCTGACAGGCCAG 1212
QY 1141 CAGAGATTGCTGACAGCAACAGCTGAGCACTGTGGGCCCGGAGAGCTCCGCTGTGCTG 1200
DB 1213 CAGAGATTGCTGACAGCAACAGCTGAGCACTGTGGGCCCGGAGAGCTCCGCTGTGCTG 1272
QY 1201 CTCTGACAGATGATCCGCACTTCCACGTGTCTCTGAGACAGAGAGCAAGGGGGAGAG 1260
DB 1273 CTCTGACAGATGATCCGCACTTCCACGTGTCTCTGAGACAGAGAGCAAGGGGGAGAG 1332
QY 1261 AACACCAAGCTGAGATCTCTTAAAGGCCACATCTGAGGAATCTTCCGCCCAAGTTCTG 1320
DB 1333 AACACCAAGCTGAGATCTCTTAAAGGCCACATCTGAGGAATCTTCCGCCCAAGTTCTG 1392
QY 1321 ----- 1320
DB 1393 CTCCCTCAACCGCTGAGCTCAATCCGAGAGTGCAGAAACCCCTGATGACAGCTGTGG 1452
QY 1321 ----- 1320
DB 1453 TACACAGGGGCATCCCGAGGGCAGAGTGGCTGAGCTGTGTCATCTGGGGACTTTC 1512
QY 1321 ----- 1320
DB 1513 CTGTGTGGGAGAGCCAGGGCAAGAGATACGTGTCTGTGTGTGTGTGTGTGTGTGTGT 1572
QY 1321 -----AACCTGTACCGACTGTGAAGGGAGGCTTT 1350
DB 1573 CCCCCGACTTGCATCATCAAGTCTTGTGANTAACTGTACGACTGGAAGGGAGGCTTT 1632
QY 1351 CCTAGATTTCTTTGTGCTATGACCACTTGTAGACCCAGAGCCTTACCAAGAG 1410
DB 1633 CCTAGATTTCTTTGTGCTATGACCACTTGTAGACCCAGAGCCTTACCAAGAG 1692
QY 1411 AGTGTGTGTGTCTGTGACAGGGCTGTGCCCCAAGAGCAATGTGTGTGTGAACCATGAGGAC 1470
DB 1693 AGTGTGTGTGTCTGTGACAGGGCTGTGCCCCAAGAGCAATGTGTGTGTGAACCATGAGGAC 1752

QY	481	AAAGCAAGGACCGGACAAAGCCAAAGACAAGATATGGGGACAGCTGTGAAACTCTTT	540
Db	556	AAAGCAAGGACCGTACAAAGCTTAAGACAAGTATGTGCCAGCTGTGGAAGCTCTTT	615
QY	541	GCTGACCAACAACCGCTATGTGTGTGGGCTGTGGCTGTGGCAGCTTACAACAAGACAACAC	600
Db	616	GCTCAACCAACAACCGTATGTGTGTGGGCGTGGGGGTGTGGCAGCTTACAACAACAACAAC	675
QY	601	CACCAAGCTCTGTGCTGCCCGGCTGTGTGGGTCACTGCAGAGACCTTGCACAGAGAGATGGCT	660
Db	676	CACCAAGCTCTGTGCTGCCCGGCTGTGTGGGTCACTGCAGAGACCTTGCACAGAGATGGCT	735
QY	661	TGCAATCCGAAAGAGATCTGTGAGAAATCACTGGAGATTACACGCTGTGTGACAGATAGAG	720
Db	736	TGCAATCCGAAAGAGATCTGTGAGAAATCACTGGAGATTACACGCTGTGTGACAGATAGAG	795
QY	721	GTGTGTGGCATTCAACCGGAGATGGCTCAGCTGTGCCCGCATTCAGACTGTGAGGCTTAG	780
Db	796	GTGTGTGGCATTCAACCGGAGATGGCTCAGCTGTGCCCGCATTCAGACTGTGAGGCTTAG	855
QY	781	TACCAAGGCTTCTGTGCAAGATATGGATCGGCACTGACGTGCCACCTGTGTCAAGTTCC	840
Db	856	TACCAAGGCTTCTGTGCAAGATATGGATCGGCACTGACGTGCCACCTGTGTCAAGTTCC	915
QY	841	GATGATGTACCTGTGTTGAGAGAGGTGCAACCCGTGAGCCTGTGGGAGCTTCCACTGAACAAG	900
Db	916	GATGATGTACCTGTGTTGAGAGAGGTGMAACCGCTGAGCCTGTGGGAGCTTCCACTGAACAAG	975
QY	901	CTGACTGTGAGAGAGGTGTGACACAACGCTGACCTTCAAGTACAGATATGAGCTGTGTGGCC	960
Db	976	CTGACTGTGAGAGAGGTGTGACACAACGCTGACCTTCAAGTACAGATATGAGCTGTGTGGCC	1035
QY	961	ACCGAGATGTGTTCAGGGCGGCAAGATATGGTTACGACAGCTGCAACAGAGACTTCGGAT	1020
Db	1036	ACCGAGATGTGTTCAGGGCGGCAAGATATGGTTACGACAGCTGCAACAGAGACTTCGGAT	1095
QY	1021	GAAGAAGGAACAACCCACCCCGGAGGGGTGCAGCTGTGGGAGAAAGGCAAGTGTG	1080
Db	1096	GAAGAAGGAACAACCCACCCCGGAGGGGTGCAGCTGTGGGAGAAAGGCAAGTGTG	1155
QY	1081	CAAGAAGCACTGCAGGGGCTCAGGTAGCGCTGTGTGACAGCCAGGCCAAGCTCAGGGCCAG	1140
Db	1156	CAAGAAGCACTGCAGGGGCTCAGGTAGCGCTGTGTGACAGCCAGGCCAAGCTCAGGGCCAG	1215
QY	1141	CAGAGATTGTGCAGAACCAAGCTGAGAGACCTGTGGCCCCGCGAGCCCCGCTGTGTG	1200
Db	1216	CAGAGATTGTGCAGAACCAAGCTGAGAGACCTGTGGCCCCGCGAGCCCCGCTGTGTG	1275
QY	1201	CTCTGTGCAAGATATGACCCGCCACTCCACCTGTGTCTGTGAGACAGAGACGAGAGGGGGAAG	1260
Db	1276	CTCTGTGCAAGATATGACCCGCCCACTCCACCTGTGTCTGTGAGACAGAGACGAGAGGGGGAAG	1335
QY	1261	ACACCCACGCTGAGAGATCTTTAAGAGCCACAATCTCAGGAAATCTTCCGCCCCAAGTTCTCG	1320
Db	1336	ACACCCACGCTGAGAGATCTTTAAGAGCCACAATCTCAGGAAATCTTCCGCCCCAAGTTCTCG	1395
QY	1321	-----	1320
Db	1396	CTCCCTCAACGGCTGACAGCTATTCCGAGAGGTGCAAGGCCCCCTGATGAGCAGCTGTG	1455
QY	1321	-----	1320
Db	1456	TACCAAGGAGCCATCCGAGAGGACAGAGTGGCTGAGCTGTGTGTGCACTTGTGGGACTTC	1515
QY	1321	-----	1320
Db	1516	CTGGTCCGGAGAGCCAGGGGCAAGCAGAGTATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	1575
QY	1321	-----AACCTGTACCGACTTGAAGGGGAAGCTTT	1350
Db	1576	CCCCGGCACTTCATCATCAAGTCTTTGGATATACCTGTATCCGACTGTGAAGGGGAAGCTTT	1635
QY	1351	CTTACGATCTCTTTTGTCTATGACCACTTACTGAGACCCACAGACGCCCTTCAACAAGAG	1410

Db	1636	CTTAGACTTCCTTTGCTGCTCATGTACCACTTACTAGACCCAGACGCCCTTACCAAGAG	1695
QY	1411	AGTGTGTTGTCCTGCA.CA.GGGCTGTGTCCCAAGGACA.GTGGGTCTGAACCATGAGAC	1470
Db	1696	AGTGTGTGTGTCCTGCA.CA.GGGCTGTGTCCCAAGGACA.GA.GTGGGTCTGAACCATGAGAC	1755
QY	1471	CTGTGTGTGGGTGAGAGATTGGA.CGGGGGAA.CTTTGGCGAAGTTCAGGCGACGCTG	1530
Db	1756	CTGTGTGTGGGTGAGAGATTGGA.CGGGGGAA.CTTTGGCGAAGTTCAGGCGACGCTG	1815
QY	1531	CGAGCCGACA.CA.CCTGTGTGGCGGTAA.GTCTTGTGAGAGACGCTTCCACCTGACCTC	1590
Db	1816	CGAGCCGACA.CA.CCTGTGTGGCGGTAA.GTCTTGTGAGAGAGCGCTTCCACCTGACCTC	1875
QY	1591	AAGGCCAAGTTTCTACAGGAAGCGAGGATCTCTAAGCAGTACAGCCACCCCAATGCTGT	1650
Db	1876	AAGGCCAAGTTTCTACAGGAAGCGAGGATCTCTAAGCAGTACAGCCACCCCAATGCTGT	1935
QY	1651	CGTCTCAATGGGTGTCTGCA.CC.CAGAAACAGCCCACTTACATCGTCATGAGACTGTGCA	1710
Db	1936	CGTCTCAATGGGTGTCTGCA.CC.CAGAAACAGCCCACTTACATCGTCATGAGACTGTGCA	1995
QY	1711	GGGGGCGA.CTTCTCTGACCTTCTCTCCGACCGAGGGGGGCCGCTGTGGGGTGAACCTCTG	1770
Db	1996	GGGGGCGA.CTTCTCTGACCTTCTCTCCGACCGAGGGGGGCCGCTGTGGGGTGAACCTCTG	2055
QY	1771	CTGCAGATGGTGGGGGATGCA.GCTGCTGGCATGGAATGACCTGGAGACCAATGCTGCATC	1830
Db	2056	CTGCAGATGGTGGGGGATGCA.GCTGCTGGCATGGAATGACCTGGAGAGCAATGCTGCATC	2115
QY	1831	CACCGGGA.CCTGCGCTGCTCGGA.CT.GCCTGGTGA.CAGAGA.AAATGTCTTGAAGATAGT	1890
Db	2116	CACCGGGA.CCTGCGCTGCTCGGA.CT.GCCTGGTGA.CAGAGA.AAATGTCTTGAAGATAGT	2175
QY	1891	GACTTTGGGATGTCTCCGAGAGAAAGCCGATGGGTCTATGCA.CCTCAGGGGGCTCAGA	1950
Db	2176	GACTTTGGGATGTCTCCGAGAGAAAGCCGATGGGTCTATGCA.CCTCAGGGGGCTCAGA	2235
QY	1951	CAAGTCCCCGAGAGTGGAA.CCGCA.CT.GAGGGCCCTTAACTPACGGGCGGTACTCTCCGAA	2010
Db	2236	CAAGTCCCCGAGAGTGGAA.CCGCA.CT.GAGGGCCCTTAACTPACGGGCGGTACTCTCCGAA	2295
QY	2011	AGCAGCTGTGAGACTTTGGCATCTTGTCTGTGGAGA.CTTTCAG.CTGTGGGGGCTTCCCC	2070
Db	2296	AGCAGCTGTGAGACTTTGGCATCTTGTCTGTGGAGACCTTTCAG.CTGTGGGGGCTTCCCC	2355
QY	2071	TATCCCAACTCAGCAATCAGCAGACA.CAGGAGTTTGTGGAGAAAGGGGGGCGTCTGCC	2130
Db	2356	TATCCCAACTCAGCAATCAGCAGACA.CAGGAGTTTGTGGAGAAAGGGGGGCGTCTGCC	2415
QY	2131	TGCCCAGAGCTGTGTCCGATGTCGCTGTTCAGGCTCATGAGACAGTGTGGGCTTATGAG	2190
Db	2416	TGCCCAGAGCTGTGTCCGATGTCGCTGTTCAGGCTCATGAGACAGTGTGGGCTTATGAG	2475
QY	2191	CCTGGGACCGGACCACTTCAGCAGCACTTACAGGAGACTGCAAGCATCCGAAAGCGG	2250
Db	2476	CCTGGGACCGGACCACTTCAGCAGCACTTACAGGAGACTGCAAGCATCCGAAAGCGG	2535
QY	2251	CATCGG 2256	
Db	2536	CATCGG 2541	

RESULT 6					
AX281847	AX281847	2889 bp	DNA	linear	PAT 02-NOV-2001
LOCUS					
DEFINITION	Sequence 256 from Patent WO0177389.				
ACCESSION	AX281847				
VERSION	AX281847.1	GI:1660998			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source

1 Shiffman, D., Somogyi, R., Lawn, R., Sellhame, J.J., Porter, G.J., Miki, T. and Tai, J.
Genes expressed in foam cell differentiation
Patent: WO 017389-A 256 18-OCT-2001;
Incyte Genomics, Inc. (US)
Location/Qualifiers
1. 2889
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 997347.6"

ORIGIN

Query Match 89.7%; Score 2023.4; DB 6; Length 2889;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 2255; Conservative 0; Mismatches 1; Indels 211; Gaps 2;

QY 1 ATGGGCTTCTTCTGAGCTGTGACGCCCCCAGGGCCACGGGGTCTCTGACGCAATGCGAG 60
DB 188 ATGGGCTTCTTCTGAGCTGTGACGCCCCCAGGGCCACGGGGTCTCTGACGCAATGCGAG 247
QY 61 GAGGCGAGCTTGTCTACTGAGGGCATGAGAAAGTGGCCAGGGGCTCAAGAT 120
DB 248 GAGGCGAGCTTGTCTACTGAGGGCATGAGAAAGTGGCCAGGGGCTCAAGAT 307
QY 121 GAGGCGAGTATGACGAGCTTCTCAACATGTCCCTGACGAGCAGTGGGGCCAGAGC 180
DB 308 GACAGGAGATGACAGAGCTTCTCAACATGTCCCTGACGAGCAGTGGGGCCAGAGC 367
QY 181 CGGGCCATCAGCCCTGACAGCCCCCATCATGCTCTGGGCTGAGATCACAGCCAACT 240
DB 368 CGGGCCATCAGCCCTGACAGCCCCCATCATGCTCTGGGCTGAGATCACAGCCAACT 427
QY 241 GA-GGGCTGAGCGGCTGTGCGGAGACGCGAGAGATCTGAATCAAGGGCCCTGAG 299
DB 428 GAGGGCTGAGCGGCTGTGCGGAGACGCGAGAGATCTGAATCAAGGGCCCTGAG 487
QY 300 CAGCTGAGCTGTCTCATCTGGAGACGGGACGCTTCCGACAGCCTACAGCGAGAGTG 359
DB 488 CAGCTGAGCTGTCTCATCTGGAGACGGGACGCTTCCGACAGCCTACAGCGAGAGTG 547
QY 360 GCGAGCTGTGACGAGAGCTTCAACAGCCACAGCGAGACATTTGAGAGCTTGAAG 419
DB 548 GCGAGCTGTGACGAGAGCTTCAACAGCCACAGCGAGACATTTGAGAGCTTGAAG 607
QY 420 CAGGTACGAGCTGTGACGAGAGCTTCCGACAGCGAGACATTTGAGAGCTTGAAG 479
DB 608 CAGGTACGAGCTGTGACGAGAGCTTCCGACAGCGAGACATTTGAGAGCTTGAAG 667
QY 480 CAAAGCAAGAGACCTGTGACAGAGCCAGAGCAAGTATGTGCGAGCTGTGAGAGCTTT 539
DB 668 CAAAGCAAGAGACCTGTGACAGAGCCAGAGCAAGTATGTGCGAGCTGTGAGAGCTTT 727
QY 540 TGTCTACCAACACGCTATGTGTGTGCGGCTGTGCGGCTGTGCGAGCTTACCAACAGCA 599
DB 728 TGTCTACCAACACGCTATGTGTGTGCGGCTGTGCGGCTGTGCGAGCTTACCAACAGCA 787
QY 600 CCAACGAGCTTCTGTGCTGCGGCTGTGCGGCTGTGCGGCTGTGCGAGCTTACCAACAGCA 659
DB 788 CCAACGAGCTTCTGTGCTGCGGCTGTGCGGCTGTGCGGCTGTGCGAGCTTACCAACAGCA 847
QY 660 TTGCACTCTGAGAGAGATCTGTGAGAGATTAAGAGATTTAGAGCTGTGAGAGATTA 719
DB 848 TTGCACTCTGAGAGAGATCTGTGAGAGATTAAGAGATTTAGAGCTGTGAGAGATTA 967
QY 720 GGTGTGGGCTTACCGGAGAGATGTGTGAGCTGTGCGGCTGTGCGAGCTTACCAACAGCA 779
DB 908 GGTGTGGGCTTACCGGAGAGATGTGTGAGCTGTGCGGCTGTGCGAGCTTACCAACAGCA 967
QY 780 GTACCAAGGCTTCTGTGAGAGATGTGTGAGCTTACCAACAGCTTACCAACAGCTT 839

DB 968 GTACCAAGGCTTCTGTGAGAGATGTGTGAGCTTACCAACAGCTTACCAACAGCTT 1027
QY 840 CCAATGATCATCTGTGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTCAAGTGAACGA 899
DB 1028 CCAATGATCATCTGTGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTCAAGTGAACGA 1087
QY 900 GCTGACTGTGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTGGAGAGCTTGGAG 959
DB 1088 GCTGACTGTGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTGGAGAGCTTGGAG 1147
QY 960 CACCGAGATGTGTGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTGGAGAGCTTGGAG 1019
DB 1148 CACCGAGATGTGTGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTGGAGAGCTTGGAG 1207
QY 1020 TGAAGAGAGAACACCCACCCCGGAGAGCTGTGAGAGCTTGGAGAGCTTGGAGAGCTT 1079
DB 1208 TGAAGAGAGAACACCCACCCCGGAGAGCTGTGAGAGCTTGGAGAGCTTGGAGAGCTT 1267
QY 1080 GCAAGAGAGAGAGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTGGAGAGCTTGGAG 1139
DB 1268 GCAAGAGAGAGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTGGAGAGCTTGGAG 1327
QY 1140 GCAAGAGAGAGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTGGAGAGCTTGGAGAG 1199
DB 1328 GCAAGAGAGAGAGAGAGTGAACCGCTGAGAGCTTGGAGAGCTTGGAGAGCTTGGAGAG 1387
QY 1200 GCTCTGAGAGATGACCGGCACTTCAAGCTGTGCTGAGAGAGAGAGAGAGAGAGAG 1259
DB 1388 GCTCTGAGAGATGACCGGCACTTCAAGCTGTGCTGAGAGAGAGAGAGAGAGAGAGAG 1447
QY 1260 GACACCCAGAGAGAGATCTTTAAGAGCCATCTTCAAGAACTTCCGCCCAAGTTCTC 1319
DB 1448 GACACCCAGAGAGAGATCTTTAAGAGCCATCTTCAAGAACTTCCGCCCAAGTTCTC 1507
QY 1320 G----- 1320
DB 1508 GCTCCCTCAACGCGCTGAGACTATTCGAGAGTGAAGAGCCCTGATGAGAGCTGTG 1567
QY 1321 ----- 1320
DB 1568 GTACCAAGGAGCATCCGAGGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1627
QY 1321 ----- 1320
DB 1628 CCTGTGCGGAGAGAGAGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1687
QY 1321 ----- 1320
DB 1688 GCGCGGAGCATTCATCATCATCATCTTGTGATTAACCTGTACCAAGTGAAGAGAGAGCTT 1747
QY 1350 TCTTAGCATCTTTTGTCTGATGACCACTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1409
DB 1748 TCTTAGCATCTTTTGTCTGATGACCACTTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1807
QY 1410 GAGTGTGTGTGTCTGAG 1469
DB 1808 GAGTGTGTGTGTCTGAG 1867
QY 1470 CCTGTGTGTGTGTGAG 1529
DB 1868 CCTGTGTGTGTGTGAG 1927
QY 1530 GCGAGCGAGCAACACCTGT 1589
DB 1928 GCGAGCGAGCAACACCTGT 1987
QY 1590 CAAAGGCAAGTTTCTTCAAG 1649
DB 1988 CAAAGGCAAGTTTCTTCAAG 2047
QY 1650 GGTCTCATTTGA 1709

2048 GCGTCTCATTTGATGTCTGACCCAGAGAGCCCATCTACATGTCATGAGCTTGTGCA 2107
QY 1710 GGGGGGGGAGCTTCTTACCTTCTCGCAACGAGGGGGGCGCTCGGGGTGAAGACTCT 1769
DB 2108 GGGGGGGGAGCTTCTTACCTTCTCGCAACGAGGGGGGCGCTCGGGGTGAAGACTCT 2167
QY 1770 GCTGAGATGTTGGGGAGATGACCTGTCATGAGTACCTGAGAGCAAGTGTCTGCA 1829
DB 2168 GCTGAGATGTTGGGGAGATGACCTGTCATGAGTACCTGAGAGCAAGTGTCTGCA 2227
QY 1830 CCACCGGAGACTGCTGCTGCTGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1889
DB 2228 CCACCGGAGACTGCTGCTGCTGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2287
QY 1890 TGACTTTGGGATGCTCCGAGAGAGAGCCGATGAGGCTCTATGCAAGCTCAGGGGCTCAG 1949
DB 2288 TGACTTTGGGATGCTCCGAGAGAGAGCCGATGAGGCTCTATGCAAGCTCAGGGGCTCAG 2347
QY 1950 ACAAGTCCCGGTGAGTGAAGCCGACCTGAGCCCTTAACCTACGCGCTTACTCTCCGA 2009
DB 2348 ACAAGTCCCGGTGAGTGAAGCCGACCTGAGCCCTTAACCTACGCGCTTACTCTCCGA 2407
QY 2010 AACGACGTGTGAGCTTGTGCACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2069
DB 2408 AACGACGTGTGAGCTTGTGCACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2467
QY 2070 CTATCCCACTGAGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2129
DB 2468 CTATCCCACTGAGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2527
QY 2130 CTGCCCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2189
DB 2528 CTGCCCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2587
QY 2190 GCTCGGAG 2249
DB 2588 GCTCGGAG 2647
QY 2250 GCATCGG 2256
DB 2648 GCATCGG 2654

RESULT 7
MPFESCR 2680 bp mRNA linear ROD 05-JUN-1995
DEFINITION Mouse c-fes proto-oncogene mRNA for c-fes protein.
ACCESSION X12616
VERSION X12616.1 GI:50955
KEYWORDS fes cellular oncogene; fes proto-oncogene; proto-oncogene; tyrosine kinase.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2680)
WILKE, A.F. and KURBAN, R.R.
TITLE Isolation and structural analysis of murine c-fes cDNA clones
JOURNAL Oncogene 3 (3), 289-294 (1988)
MEDLINE 89083198
PUBMED 3060793

FEATURES
source
location/Qualifiers
1..2680
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_id="lambda gcl1."
60..2522
/note="unmated protein product; c-fes protein"
/codon_start=1
/protein_id="CAA1138.1"
/db_xref="GI:50956"
/db_xref="GOA:P16879"

/db_xref="UniProt/Swiss-Prot:P16879"
AGLHMSLSDSGGSGSPSPVSGVAGVQOMQAEHLLEGKMKMAQRYVSDREY
AGLHMSLSDSGGSGSPSPVSGVAGVQOMQAEHLLEGKMKMAQRYVSDREY
SVLREHRSIRKTVNEMOQLODELTKHSDIIEKTKOYRLTVDSQARRKYDAS
KQKDXDKADKTVRSIMKLPANRNVYLVGVAQALJHHHHNHPMLPGLOSLQDHEE
MAGLIKDLIOEYLEISSLVQDDVASIHRLEAAALAPRLOPEFYLGRQYGTDPVP
CVTFPESLLEDESOLEPGELOINLETLSVQHTLVSDDELAVARKVEYLSREMSQL
QRELSERQNTHPREPVOLSKRMLOEALQELQIALSGODLQAOEHLQSKMBOG
TGEPAVPLLOPDRSTSTEREGRTPLLEIKSHFSGI PPKPSI PEPLOLQVPEQ
KPLVPAVPLHGAIPMAEVAEILTHRGDPLVRSOQKQEVLSVMWDGHRHFIQSLD
NLYRLEGSPESIPLILTHLSSQDPLTKSSVYLPANPKDKMLKHEIDLQSGOIC
KQNFSEVSGRLRNDTPVAVASCEBTELPDLPKAFLOARLILQYHNPYRLQVGC
TQKQPIYIVMELVQGDPLFLRTGARGRLVITLLQMDGDAAGMEYLESKCIHRDL
AARNCILVTEKNYLKISDPGMSREADGIVAAQSGROYVVKMTABEALNMGYSBSB
VNSFILLMETSLASPPYPLNTNQTREFEVKGRLPCEPLCPDAVFLMGCWAYE
PGQRSPSFIQOEHLISIKRHR"

ORIGIN
Query Match 68.3%; Score 1540.4; DB 10; Length 2680;
Best Local Similarity 79.2%; Pred. No. 2,5e-258;
Matches 1954; Conservative 0; Mismatches 296; Indels 216; Gaps 2;

QY 1 ATGGGCTTCTTCTGAGCTGTGACAGCCCGGAGGCTCCTGCAAGCAATGCAAG 60
DB 60 ATGGGCTTCTTCTTCAAGCTGTGACAGCCCGGAGGCTCCTGCAAGCAATGCAAG 119
QY 61 GAGGCGGAGCTTCTTCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
DB 120 GAGGCGGAGCTTCTTCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179
QY 121 GACAGGAGATGATGACAGAGCTGTTCAACAATGCTCCTGACAGAGAGTGGGGGCGAGAC 180
DB 180 GACCGGAGATGATGACAGAGTGTCTTCAACAATGCTCCTGACAGAGAGTGGGGGCGAGAC 239
QY 181 CGGCGCATAGAGCCCTGACAGCCCGGAGGCTCCTGAGGCTGAGATCAACAGCCAACT 240
DB 240 TGGAGAGAGGAGGCGCGGAGAGGCTCCTGAGGCTGAGATCAACAGCCAACT 299
QY 241 GAGGCGCTGAGGCGGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 300 GAGGCGCTGAGGCGGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 359
QY 301 AAGCTGAGCTGCTCATCCGGGAAACGGAGAGAGCTTCCGAAAGCTTACAGAGAGAGAGAG 360
DB 360 AAGCTGAGCTGCTCATCCGGGAAACGGAGAGAGAGCTTCCGAAAGCTTACAGAGAGAGAGAG 419
QY 361 CAGCAGCTCAG 420
DB 420 CAGCAGCTCAG 479
QY 421 CAGTACGAGCTTGTGCAAGGAG 480
DB 480 CAGTACGAGCTTGTGCAAGGAG 539
QY 481 AAAG 540
DB 540 AAAG 599
QY 541 GCTCAGCAG 600
DB 600 GCTCAGCAG 659
QY 601 CACGAGCTCTGCTGCGCGGCTGCTGCGGAGTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 660 CACGAGCTCTGCTGCGCGGCTGCTGCGGAGTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
QY 661 TGCATCTGAAGAGATCTGCAAGAAATCTGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 720 TGCATCTGAAGAGATCTGCAAGAAATCTGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 779
QY 721 GTGTGCGCATTCACCGGAG 780
DB 780 GTGTGCGCATTCACCGGAG 839

QY 781 TACCAAGGCTTCTGGAAGTAATGGTCCGCACTGACCTCCACCTGTGTCAAGTTT 840
DB 840 TACCTAGGCTTCTGGAAGTAATGGTCCGCACTGACCTCCACCTGTGTCAAGTTT 899
QY 841 GATAGTCACTGTTGAGAGGAGTGAACGCTGAGAGCTGGAGAGCTCCAGTGAAGAG 900
DB 900 GATAGTCCGTTTGTGAAGCGGGGAACAGCTGAGAGAGAGAGAGAGAGAGAGAG 959
QY 901 CTGACTGTGAGAGGCTGAGCAACGCTGACCTCAGTGAAGATGAGTGTGTGAGCC 960
DB 960 TTGACCTTGGAGAGCGTGAACACAGCTGACTTGTGACAGATGAACTGCTGTGAGCC 1019
QY 961 ACCGAGATGCTGTTCAGGCGGAGAGATGTTACGAGCTGCAACAGAGAGCTCCGAAAT 1020
DB 1020 ACCAAGAGGTGTGAGCGGAGAGATGTTACGAGCTGAGCTGAGCTGCAAAAGT 1079
QY 1021 GAAGAGAGAGAACCCACCCCGGAGCGGAGTCAAGCTGTGGCAAGAGAGAGTGTG 1080
DB 1080 GAGGAACAGAACCCACCCCGGAGCGGAGTCAAGCTTGTGAAGAGAGAGAGATGCTG 1139
QY 1081 CAAGAGACATGAGAGGAGCTGAGAGTACGCTGTGACAGGAGCAAGCTGACAGCCAG 1140
DB 1140 CAAGAGACATGAGAGGAGTGAAGATAGCACTGTGACAGGAGCAAGCTTCAAGCCAG 1199
QY 1141 CAGGAGTGTGAGAGCAAGCTGAGAGCACTGGGCGCCGAGAGCCCGCTGTGTGCTG 1200
DB 1200 CAGGAGCTTGTGAGAGCAAGATGAGAGCACTGGGCGAGGCGAGCCCGCTGTGTGCTG 1259
QY 1201 CTCTCTGAGATGAGCGGAGCTGAGCTGCTGCGAGAGAGAGAGAGAGAGAGAG 1260
DB 1260 CTCTCTGAGATGAGCGGAGCTGAGCTGCTGCGAGAGAGAGAGAGAGAGAGAG 1313
QY 1261 ACAAGGAGCTGAGAGTCTTAAAGAGCAATCTCAGAAATCTTCCGCGCAAGTTCTC- 1319
DB 1314 ACAAGGAGCTGAGAGTCTTAAAGAGCAATCTTCCGAAATCTTCCGCGCAAGTTCTC- 1373
QY 1320 ----- 1319
DB 1374 ATCCCTCAAGCTGAGCTGTGCGGAGAGTGAAGAGCTCTCTATGAGAGCTGTG 1433
QY 1320 ----- 1319
DB 1434 TACCAAGGAGCTATCCCTGAGAGAGTGTGAGCTGTCTAAAGCACTGGAGACTTC 1493
QY 1320 ----- 1319
DB 1494 CTGAGTTCGAGAGAGCAAGGAGAGAGATATGTAAGTGTGATGTGAGATGAGCAG 1553
QY 1320 -----GAACTGTATCCGAGCTGAGAGAGAGAGAGAGAGAGAGAG 1553
DB 1554 CCGGAGATTTATCATCATTCAGTCTTATGATTAAGTGAAGAGAGAGAGAGAGAG 1613
QY 1351 CCGTATGCTTGTGATGAGCACTGAGCACTGAGAGAGAGAGAGAGAGAGAGAG 1613
DB 1614 CCGAGATGATCTGTGATGAGCACTGAGCACTGAGAGAGAGAGAGAGAGAGAG 1673
QY 1411 AGTGTGTGTCTGAGCAAGGAGCTGTGAGCAAGAGAGAGAGAGAGAGAGAGAG 1470
DB 1674 AGTGTGTGTCTGAGCAAGGAGCTGTGAGCAAGAGAGAGAGAGAGAGAGAGAG 1733
QY 1471 CTGAGTGTGTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1530
DB 1734 CTGAGTGTGTGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1793
QY 1531 CGAGCGGAG 1590
DB 1794 CGTGTGAG 1853
QY 1591 AAGGAG 1650
DB 1854 AAGGAG 1913

QY 1651 CGTGTATGTTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1710
DB 1914 GGTGTATGAG 1973
QY 1711 GGGGAG 1770
DB 1974 GGGGAG 2033
QY 1771 CTGAGATGAG 1830
DB 2034 CTGAGATGAG 2093
QY 1831 CAGCGGAG 1890
DB 2094 CAGCGGAG 2153
QY 1891 GACTTGTGAG 1950
DB 2154 GACTTGTGAG 2213
QY 1951 CAGTGTGAG 2010
DB 2214 CAGTGTGAG 2273
QY 2011 AGGAG 2070
DB 2274 AGGAG 2233
QY 2071 TATCCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2130
DB 2334 TATCCCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2293
QY 2131 TGCCAG 2190
DB 2394 TGCCAG 2453
QY 2191 CCGGAG 2250
DB 2454 CCGGAG 2513
QY 2251 CATCGG 2256
DB 2514 CATCGG 2519

RESULT 8
ACF7S140A
LOCUS
DEFINITION
ACF7S140A 2715 bp ss-RNA 1linear VRL 27-APR-1993
Fujinami sarcoma virus temperature sensitive (ts) pl40 transforming
protein RNA, 3' end.
ACCESSION
M14930
VERSION
M14930.1 GI:209688
KEYWORDS
transforming protein.
SOURCE
Fujinami sarcoma virus
ORGANISM
Fujinami sarcoma virus
REFERENCE
1 (bases 1 to 2715)
Chen, L.H., Haseida, E., Wheatley, W., and Lee, W.H.
AUTHORS
Single amino acid substitution, from Glu1025 to Asp, of the
TITLE
oncogenic protein causes temperature sensitivity in transformation
and kinase activity
JOURNAL
Virology 155 (1), 106-119 (1986)
MEDLINE
87044080
PUBMED
2877522
COMMENT
Original source text: Fujinami sarcoma integrated viral DNA, clone
pFL-5 from ts FL-15.
FEATURES
source
1..2715
/organism="Fujinami sarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11885"
CDS
1..2697
/note="pl40 transforming protein"

```
/codon_start=1
/protein_id="AAA42403.1"
/db_xref="GI:209689"
/translation="PHDVTNPMRVILGPAFYALIMPMMAASQIHPPOEHTSTSA
AGTWRLTQASBSRHRLPHCSAPSHODHSMGFPEPLWCPKSGISELLRLQDSLELLE
LMKRMNSORAKSDREYAGMLHMFSQLERKDEGLHRLADHSDHWMVLSQET
LSQTLRRHAEELAPLAKSLIRDOQIARKVSEBOMQLOSEYMTQOEWKEXKA
QYBSLVRDSTQAKRKYOASIKEREKAEKVKYSLKYLALNVOAVLVAQAAHHR
HYORALPTLHESLYLODMVLTKEIIGVCSITSLVQEDVLTIOEVAHVEHD
PATYBSFYQCHRYDSEVPVAVTFLSEELBASLEPQLQNELTIESVQHSITE
BELLSARKAVSKEQVWELQVLRGBELSLSEBVLHUKRKQREBQOQIQSLVC
AQAQLQADRMPLTNKLAEGLSDPEPALPQEDROASCTDORSVTLAKTINHS
GIFSPRSLPEVPPLIPEVQKPLCOQAMYGAI PRSEVQELLYDSFVLRSEQKOE
YVLSVMDQOPRHFIIOADNLYRLBDGPTIPLIDHLOSQRPIYRSQSVLTPRA
VLKDVLMHEDVILGRIIGRNKGVSEFSGRLADNTPVAVNSCRETLRPELAKRELQ
EARIKQCHNPVIRLIVCTQKQPIYIYVWELVQSDPLFSRLSKPRLMKKLITOM
DNAAMGEYIESHCHTIRDLAANRCIVTEKNTKISDFGSRBEBGVASTGAMKOI
PVKMTAPEALNTGWSSESDVMSFGILMEAFSLGAVPYANLSNQTRAIIBGVRLB
PREQCPEDVRLMQRCMEYDPHRRPSFGAVHODLAIIRKMR"
```

ORIGIN

Query Match 45.1%; Score 1018.2; DB 14; Length 2715;

Best Local Similarity 66.2%; Pred. No. 2.2e-167;

Matches 1641; Conservative 0; Mismatches 608; Indels 228; Gaps 3;

```
QY 1 ATGGGCTTCTCTTGAAGCTGTGAGCGCCCGGAGCCACCGGGTCTTGACGAATATGAC 60
Db 223 ATGGGCTTTGGGCGGAGCTGTGTGCTCCGAAAGGGGACAGTACGTGCTGGCTGAC 282
QY 61 GAGGCGGAGCTTGTCTACTGAGAGGCGATGAGAAAGTGATGGCCGCGGTCAAGAGT 120
Db 283 GACAGCGAGTTGGCGCTCTGAGAGCTGATGAAGAAGTGATGTCAAGCTGCGCAAGAGC 342
QY 121 GACAGGAGATATGACAGACTGCTTCAACCAATGTCC-----CTGACGACAGT 168
Db 343 GACCGGAGATACCGGGGAGTGTGACCAACATGTTCTCAGCTGAGAAACAGAGGGC 402
QY 169 GGGGGCCAGAGCGGGGCGATCAGCCCTGACAGCCCATCATGTCAGTCTGGGCTGAGATC 228
Db 403 CTGGGACATCTCCGTGCCACCGACCAAGACCAAGTCCGGGAGTGTGTGGGTTCTG 462
QY 229 ACCAGCCMACTGAGGGGCTGAGCGGCTTCTGCGGAGCAGCAGAGAGATCTGAATCA 288
Db 463 GCAAGCCAGACTGAGACCTGAGCAGACGCTGGGCGGACGCGGAGAGAGCTGGACGC 522
QY 289 GGGGCTTGAAGCAAGCTGAGCTGCTCATTCGGGGAAGGACGACAGCTTCCCAAGACTTAC 348
Db 523 GGGCCGCTGGCAAGCTGAGCATATGATCCGTGACAAAGACAGAGCTTCCCAAGGCTTTC 582
QY 349 AGCGAGAGTGGGAGAGCTGACAGAGAGCTCAACAAAGACCCACAGCCAGACATTGAG 408
Db 583 AGTGAAGATGGGAGAGAGCTCAGCCAGAGATACCCCTTGACACACAGCAGAGAGTTGAG 642
QY 409 AAGCTGAAGACGAGTACCGAGCTCTGGCACGGGACAGTCCCAAGCCAGCCCAAGTAC 468
Db 643 AAGCTGAAGACGAGTACCGAGCTCTGGTGTGACAGCAACCCAGGCGCAAGCCCAAGTAC 702
QY 469 CAGGAGCCGCAAAAGCAAGAGACCTGACAAAGGCCAAGACAAAGTATGAGCCGAGCTG 528
Db 703 CAGGAGCCGCAAAAGCAAGAGACCGGAAAGCGAAAGAAATATCGTGCAGAGCTG 762
QY 529 TGGAGGCTTTTGTCTCAACAACCGCTATGTGTGGGCTGCGGGCTGCGGAGCTTAC 588
Db 763 TCGAAGCTTATGTCTTGCACACACAGTACGTGTGCTGTGACAGGCGGCTGCTTGCAC 822
QY 589 CACGAGACCAACAACAAGCTCTGCTGCGCGGCTGTGCGGTCACTGACAGACTTGAC 648
Db 823 CACGACATCACTACAGCGGTGCTGTCCACCTTGACAGAGTCCCTTCAAGCTTGAC 882
QY 649 GAGGAGTGGCTTGCATCTGGAAGAGATCTGTGAGAAATACCTGAGAGATTAGAGCTG 708
Db 883 CAGGAGTGGTCTTGTGTTGAAAGAGATTCTTGGGAGAGTACTGACATCAACAAGCTG 942
```

```
QY 709 GTGACGATGATGATGGTGCATTCAACCGGAGATGCTGACAGCTGCCCCGATCCAG 768
Db 943 GTGACGAGGAGAGCTGTGGCTTCCACCAAGAGAGTGTGCCACGCGCTCCAGATGATCGAC 1002
QY 769 CTTGAGGCTGAGTACCAAGGCTTCTTGCAAGATATGGTCCGACCTTGAAGTCCACCC 828
Db 1003 CCGGCACTGAGATACGACAGCTTGTTCAGTGCACAGGTACGACTGAGGTACACACCG 1062
QY 829 TGTGTACGTTGATGATGACATGCTGCTTGGAGAGGGTGAACCGCTGAGGCTGGAGCTC 888
Db 1063 GCAATGACCTTTGATGAGAGCTTGTCTGGAGAGGACAGAGAGCTTGAAGCCAGGAGAGCTG 1122
QY 889 CAGCTGAAGAGCTGATCTGTGAGAGAGCTTGACACACGCTGACCTCACTGACAGATGAG 948
Db 1123 CAGCTGAATGAGCTGATCCATCGAGAGCGTGCAGACATCCCTGACATTCATTAAGAGGAG 1182
QY 949 CTGGCTGTGGCCACCGAGATGGTGTTCAGGCGGAGAGATGTTACGACAGCTGCAACAG 1008
Db 1183 CTGTTGGCCAGCAGAGAGGACAGTGAACAGACAGACAGCGGCTGTGGAGCTGACAGGTT 1242
QY 1009 GAGCTCCGGAATGAGAGAGAGAACACCCACCCCGGAGCGGGTGGAGCTGGGCGAAG 1068
Db 1243 GAGCTGCGGGGAGAGAGAGCTGGACATGAGCTCTGGGAGCGGGTGAATTTGCTGGGCGAAG 1302
QY 1069 AGGCAAGTCTGCAAGAGCACTGACAGGGGCTGACAGTACGCTGTGACAGCCAGCTAAG 1128
Db 1303 CGGACAGGGGCTGCGGGAGGCCAGGACGAGCTGACAGGGCTTGGTCTGTGCTCAGGCGAAG 1362
QY 1129 CTGACAGGCCACAGAGATTTGCTGACAGACCAAGCTGGAACACTTGGGCCCCGCGAGCC 1188
Db 1363 CTGACAGGCAACAGGGATATGTCTGACCAACAACTGGGAGAGTGGGCTTGAAGAGCCC 1422
QY 1189 CCGCTGTGCTGCTGCTCTGAGAGATGACCGGCACTCCAGCTGCTCTGAGAGCAGAGCGCA 1248
Db 1423 CTTCCGCTTGGCTCCCTGACAGAGAGACCGGACATGACGCTGCTCAACGATACAGAGCCG 1482
QY 1249 GAGGGGGGAGAGACACCAACGCTGAGATCTTAAAGGCAACATCTCAGGAATCTTCCGC 1308
Db 1483 AGTGGGGTGA-----CTGCGCTGAAGACCATCAAGAACCAATCTCGGCGATCTTCAGT 1536
QY 1309 CCCAAGTTCTCG----- 1320
Db 1537 CCCAGGTTCTCGTGGCACTCCCGGCTCCCTCATCCGAGAGTGAAGAGCCGCTGTGC 1596
QY 1321 ----- 1320
Db 1597 CAGCAGGCTGTGACCAAGGGGCGCATCCCGGCTCGAGAGTGAAGAGCTGTGAATAC 1656
QY 1657 AGCGAGACTTCTGTGTGGGAGAGACAGGGGAAAGAGATATGTGCTCAGCGTGTG 1716
Db 1321 -----AACTGTATCCCACTGGA 1338
QY 1717 TGGGATGGGACGCGCGGACCTTCAATCATTCAGGCTGTGAACAACCTGTACAGGCTGGA 1776
Db 1339 GGGGAAGGCTTCTGATGACTTCTTGTGCTATGACCAACTACTGAGACCCAGACGCC 1398
QY 1777 GACGATGGCTTCCACATCCGCTGCTCATGACCACTGTGTGAGAGCCAGCCGCGCC 1836
Db 1399 CTCACCAAGAGAGTGTGTCTGTCACAGGGCTGTGCCAAGAGCAAGTGGTGTG 1458
QY 1837 ATCACCCGAGAGAGCGGACATCGTCTGACAGAGCGGTGTCAAGGACAAAGTGGTGTG 1896
Db 1459 AACCATGAGAGCTGTGTGTGGTGTGAGAGATTTGAACGGGGAACTTTGGCGAAGTCTT 1518
QY 1897 AACCATGAGAGCTGTGTGTGGGAGAGCCATTTGGCGGGGAACTTGGGAGGTGTTC 1956
Db 1519 AGCGAGCGCTGGAGCCGACAAACACCTGTGTGGCGGTGAAGTCTTGTGAGAGAGCGCTC 1578
QY 1957 AGCGGCGCTGTGTGTGACCAACCCCGTGTGGGTGAATCTGTGCGGGAAACCTGTG 2016
Db 1579 CCACTTGAACCTCAAGGCCAAGTTTCTTACAGAGAGCGAGATCTTGAAGAGTACAGCCAC 1638
```

Db 2017 CCGCCGAGCTGAGAGCCAGATTCTCTGAGAGAGCCAGATCTCAAGCAGTCCACAT 2076
Qy 1639 CCCAATCATGTGTGCTCATTTGTGTCTGACCCAGAGAGCCCATCTACATGTCTATG 1698
Db 2077 CCCAATCATGTGTGCTCATTTGTGTCTGACCCAGAGAGCCCATCTACATGTCTATG 2136
Qy 1699 GAGCTTGTGACAGGGGGGCGAATCTTCTGACCTTCTCTGACAGAGGGGGGCGGCTGCGG 1758
Db 2137 GAACTGTGACAGGGAGAGGAGCTTCTGAGCTTCTCTGACAGAGGGGGGCGGCTGAG 2196
Qy 1759 GTGAAGACTCTGTCTGAGATGTGTGGGGAGATGACAGTCTGTGACATGTGAGAGAGC 1818
Db 2197 ATGAGAGAGCTGATCAAGATGATGACATGCGCGCGGCGGCTGAGATGCTGAGAGC 2256
Qy 1819 AAGTGTGATCCAGCGGAGCCTGTCTCTGAGACTGCTGTGTGA CAGAGAGAAATGTC 1878
Db 2257 AAGCAGTGTGATCCAGCGGAGCCTGTCTCTGAGACTGCTGTGTGA CAGAGAGAAATGTC 2316
Qy 1879 CTGAAGATCAATGATCTTTGGAGATGTCCGAGAGAGAGCCGATGGGCTCTATGAGCTCA 1938
Db 2317 CTGAAGATCAATGATCTTTGGAGATGTCCGAGAGAGAGAGATGTGTCTATGCTTCAAG 2376
Qy 1939 GGGGGGCTGAG 1998
Db 2377 GGGGGGCTGAG 2436
Qy 1999 TACTCTCGAG 2058
Db 2437 TACAGCTCGAG 2496
Qy 2059 GGGGGGCTGAG 2118
Db 2497 GGGGGGCTGAG 2556
Qy 2119 GGGGGGCTGAG 2178
Db 2557 GGGGGGCTGAG 2616
Qy 2179 TGGGCTATGAG 2238
Db 2617 TGGGCTATGAG 2676
Qy 2239 ATCCGAAAGCGGAGATCG 2255
Db 2677 ATCCGAAAGCGGAGATCG 2693

RESULT 9

ACF 4788 bp ss-RNA linear VRL 27-APR-1993
LOCUS Fujinami sarcoma virus (unintegrated circular), complete genome.
DEFINITION J02194 K01827 K01828
ACCESSION J02194.1 GI:209686
VERSION 1
KEYWORDS C-myc proto-oncogene; complete genome; fps oncogene; polypeptide.
SOURCE Fujinami sarcoma virus
ORGANISM Fujinami sarcoma virus
REFERENCE 1 (bases 1 to 4788)
AUTHORS Shibuya, M. and Hanafusa, H.
TITLE Nucleotide sequence of Fujinami sarcoma virus: evolutionary relationship of its transforming gene with transforming genes of other sarcoma viruses
JOURNAL Cell 30 (3), 787-795 (1982)
MEDLINE 83050964
PUBMED 6291784
REFERENCE 2 (bases 1832 to 1881; 2851 to 2902)
AUTHORS Carlberg, K., Chamberlin, M.E. and Beemon, K.
TITLE The avian sarcoma virus PRCT1 lacks 1020 nucleotides of the fps transforming gene
JOURNAL Virology 135 (1), 157-167 (1984)
MEDLINE 84225820
PUBMED 6328746

COMMENT Original source text: FSV [Fujinami sarcoma virus], unintegrated circular DNA [1], [2]
The sequence of 1182 amino acids was deduced from the Fujinami sarcoma virus (FSV) transforming protein p130, the product of gag-fps fused gene. p130 is highly homologous to the gag-fes sequence of feline sarcoma virus (see separate entries), and is 40% homologous in the 280 residue carboxyl end (including the phosphoacceptor tyrosine residue) of Rous sarcoma virus p60 protein.

FEATURES

source 1..4788
/organism="Fujinami sarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11885"

LTR 1..347
/note="5' LTR"

repeat_region 247..267
/note="5' terminal repeat"

CDS 626..4174
/note="p130 polypeptide"
/codon_start=1
/protein_id="AAA42402.1"
/translation="MEAVIKVISAQKTYCGKTSRKEIGAMLSQLQKGLMSLSD LVSQSDWPTALTORAMVLKSGELKTMGLVLAALKAREOYVTSQAKFMVLGG GVSPPGPECTEKATERRIDKGETTVQDTCAPERTPTPTVTSCHTCGALICN CATASAPPPVYSGVPSLAVGVEQGGSGTTPKAPGVPABPGATGLAPALTDW ARIRELASTGPPVAMPVAVITGEPVAPLPKLIAGLAVAGAGLSPVAVAVE ALMSPLPHDVTPVAVILGPA PHLAMMDANAASQGLRPOPESTVSLAAGTRL TOAESRHLRPHCSAAPSVDHDSAMGEGELCKPKHTLRLRDESLTSLMCKMM SORAKSDREYALHMHPSOLKOEGLHLRATHDSOISGSMWVLSLQTELTSLRL RABELAAPLAKSLILDKOOLKRVSEQNOQSRYAANTTQDVEVLKAOYRSLV RSTQAKRYQASQKEREKAKERVSLKULALHNOYVLAIVQALHHHHTQPA LPTLHESLYSLQEEVNLVLEKILGEVCSITSLVQEBVLAHQKVAIVAMIDATYS SEVQCHRYDSVEVPVATPDESLSLEAENLEPELQINELTISVQHSLSIBELIAS RKAWSKEQVMELOVEALGEBELASPGSRVHLGROGLREAOQLOGLVCAQAKO AORDMLANKLAEISGSEPPPALPIEDRQASRSTOERSGVTAALKTIKHISGIFSR FSLPPVPIPIPEVQKPLCOQANYHAIIPSEVQELIKYSGDPIVRESQKQEVLSVL WQGPRLFTIQAADNLYRLEDGLPTIPLILHLQSQEPITRKSGIVLRAVLKRW VNHEDVHLGERIIRGNFGEVSGRLRANTVAVVSCRETPPEIKAFLOEARLIK QCNHPIVRLIVCTQKQPIYIVMEVQGDPLSPRSKGPLKMKKLKMMENAAAG MEYLSKCHIHDLAARNCLVTEKNTKLSIDGMSQEDGDGYASTGKMKQIPVKTWA PEALNYGWSSESDVWSFGILLMEAPLSIGAVPYANLSTQTRTEALRQVRLPEPCP EDVYRLMQRCHETDPRHRPSFGAHOBLIAIKRRR"

mat_peptide 629..1549
/product="gag protein"

mat_peptide 1550..4171
/product="v-fps transforming protein"

LTR 4442..4788
/note="3' LTR"

repeat_region 4688..4708
/note="3' terminal repeat"

ORIGIN 5' end of proviral genome; 768 bp upstream of SmaI site.

Query Match 45.1%; Score 1018.2; DB 14; Length 4788;
Best Local Similarity 66.2%; Pred. No. 2e-167;
Matches 1641; Conservative 0; Mismatches 608; Indels 228; Gaps 3;

Qy 1 ATGGGCTTCTCTTCTGAGCTGTGACAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 1700 ATGGGCTTCTCTTCTGAGCTGTGACAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAG 1759
Qy 61 GAGGCGAGCTTGTCTACTGAGAGGAGATGAGAAATGATGAGCCAGCGGATCAAGAT 120
Db 1760 GACAGGAGATGAGGCTCTCTGAGAGTGAAGAAATGATGACAGCGTCTAAGAGC 1819
Qy 121 GACAGGAGATGAGGAG 168
Db 1820 GACCGGAGATGAGGAG 1879
Qy 169 GGGGGCAGAGCGCGGAGATGAGCCCTGACAGCCCATGATGATCTGAGGCTGAGATC 228
Db 1880 CTGGGACATCTCTGCTCCACCGACACAGCAGCAGATCGGGAGTCTGTGTGTTCTG 1939

QY 229 ACCAGCCAACTGAGGGCCTGAGCCGCTTCTGCGGAGACCGCAGAGATCTGAACCTA 288
 DB 1940 GCAAGCCAGACTAGACACTGAGCCAGACGCTCGCGGACCGCGAGAGAGCTGGCAGCG 1999
 QY 289 GGGCCCCCTGAGCAAGCTGAGCCCTGCTCATCCGGGAAACGGCAGAGCTTGGCAAGACTAC 348
 DB 2000 GGGCCCGCTGGCCAAAGCTGAGCATATCTGATCCGTGACAAAGCAGAGCTTCCGAAGTCTTC 2059
 QY 349 AGGAGCAGTGGGAGAGAGCTGCGACAGGAGCTCAACCAAGCCCAAGCCAGGACATTAG 408
 DB 2060 AGTGAACAGTGGGAGAGAGCTCAAGCCAGGAGTACGCTTGGACCAACAGCAGGAAAGTTAG 2119
 QY 409 AAGCTGAAGAGCAGATACCGAGCTTGGCAGCGGAGACAGTCCCAAGCCAAAGCCAAAGTAC 468
 DB 2120 AAGCTGAAGGACAGATACCGGAGCTGGTGGTGGTGGACAGACCCAGGCCAAAGCTTAC 2179
 QY 469 CAGGAGGCCAGCAAGACAAAGACCGTGAACAAGGCCAAGGACAAATATGTGGCCAGCTTG 528
 DB 2180 CAGGAGGCCAGCAAGACAAAGACCGGGAAGGCGAAAGTAAAGTACGTGGCCAGCTTG 2239
 QY 529 TGGAACTCTTGTGCTACCAACACCGCTATGTGTGGGCGTGGGGGCTGGCGAGCTACAC 588
 DB 2240 TCGAAGCTCTATGCTTGGCAACAACAGTATGTGTGGCTGTGACAGGCGGCTGGCTTAC 2299
 QY 589 CACGAGACCAACCAAGCTCTGTGCTCCGAGCTGTGGCTGTGCTGTGCTGTGCTGTGCTGTG 648
 DB 2300 CACGACATATACCAAGCCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2359
 QY 649 GAGGAGATGGCTTGCATCTGTAAGAGATCTGCAAGAAATCCTGAGATTAAGCAAGCTTG 708
 DB 2360 CAGGAGATGGCTTGTGTTGTAAGAGATCTTGGGAGATCTGTAAGATCAACAGCTTG 2419
 QY 709 GTGACAGATAGAGTGGTGGCTTCAACCGGAGATGGCTGAGCTGTGCGGCAATCCAG 768
 DB 2420 GTGACAGAGAGAGCTGTGCGCATTCACAAAGAGTTGCCCCAGCTGTGAGATGATGAC 2479
 QY 769 CCTGAGGCTGATGACCAAGGCTTCTGTGCAAGATGAGTCCGACCTGAGCTGCCACC 828
 DB 2480 CCGGCCACTGATGACAGAGCTTGTGTAAGTCCAGCAAGATACATCTGTAGATACCAAG 2539
 QY 829 TGTGTACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 888
 DB 2540 GCAATGACCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2599
 QY 889 CAGCTGAACAGCTGATCTGTGAGAGGCTGACAGACCGTGAACCTTGAAGTGAAGATGAG 948
 DB 2600 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2659
 QY 949 CTGAGCTGTGACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1008
 DB 2660 CTGATGAGCCAGCAGGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2719
 QY 1009 GAGCTCCGGAATGAAAGAGAAACCCACCCCGGGAGCGGGTGCAGCTGTGGGCAAG 1068
 DB 2720 GAGCTGGGGGAGAGAGCTGTGCACTCAGCCCTGGGAGCGGGTGCAGCTGTGGGCAAG 2779
 QY 1069 AGGCAATGCTTGAAGAAAGACCTGCAAGGAGCTGAGAGTACGCTGTGACACCGCCAAAG 1128
 DB 2780 CCGGAGGAGGCTGCGGAGGCGCCAGCAGCACTGAGGAGCTTGTGTGTGCTCAGGCGCAAG 2839
 QY 1129 CTGAGGAGCCAGAGAGTGTGCAAGACCAAGCTGAGAGCACTGTGGGCCCGGAGCGCC 1188
 DB 2840 CTGAGGAGCAGAGCGGAGATATGCTGTGCGCAAGAGTGGCAGAGCTGGGCTTGAAGAGCGC 2899
 QY 1189 CCGGCTGTGTGCTCTCTGAGAGATGACCGCACTCAAGTGTGCTTCTGAGCAGAGAGCGA 1248
 DB 2900 CTTCCCGCTTGGCCCTGCAAGAGAGACCGGCAAGTCAAGCGCTTCAAGGATCAGAGCGC 2959
 QY 1249 GAGGGGGAAGACACCAAGCTGTGAGATCTTAAAGCCCAATCTTCAAGAAATCTTCCG 1308
 DB 2960 AGTGGGGTGA-----CTGGCTGAAGACCATCAAGAACCAATCTTGGGCAATCTTCAAGT 3013

QY 1309 CCCAAGTTCTCG----- 1320
 DB 3014 CCCAGGTTCTCGCTGCACTCCCGGCCCCCTCATCCGAGATGAGAAAGCGCTGTGC 3073
 QY 1321 ----- 1320
 DB 3074 CAGCAGGCTGTATACCAAGGAGCATCCGCGCTGGAGTGCAGAGAGCTGTGAATATAC 3133
 QY 1321 ----- 1320
 DB 3134 AGCGAGACTTCTGTGTGGGAGAGACGAGGGAAGCAGAGATGTGCTCAGCGTGTG 3193
 QY 1321 -----AAGCTTACCGACTGAA 1338
 DB 3194 TGGGATGGGACAGCCCGCACTTATCATCAAGCTGTGACAACTGTATACAGCTGAA 3253
 QY 1339 GGGGAAGCTTCTTATAGCATTTCTTTGTCTATGACCACTTACTGAGACCCAGCGCC 1398
 DB 3254 GACGATGGCTTCCCACTCCCGCTGTCTATGACCACTGTGTGAGAGCCAGCGGCC 3313
 QY 1399 CTCACCAAGAGATGTGTGTGCTGTCGACAGAGGCTGTGCCCAAGGACAAAGTGGTCTG 1458
 DB 3314 ATCAACCCGAGAGCGGACATCGTCTGACCAAGCCGTCTCAAGGACAAATGGATGCTC 3373
 QY 1459 AACCATGAGACCTGTGTGTGGGTGAGAGATTTGACCGGGGAACTTTGGCGAAATGTTTC 1518
 DB 3374 AACCATGAGAGAGTGTGTGTGGGGAGCGCAATTTGGCCGGGGAACTTCCGGGAGTGTTC 3433
 QY 1519 AGCGAGCGCTGTGAGAGCCGACAAACCTGTGTGGCGGTGAAGTCTGTGAGAGAGAGTTC 1578
 DB 3434 AGCGCGGCTGTGTGTGACAAACCCGTGTGGCGGTGAAGTCTGTGCGGGAAACCTTG 3493
 QY 1579 CCACTGACCTCAAGGCGCAAGTTTCTACAGAGAGCAGAGATCTGAAAGATGACAGCCAC 1638
 DB 3494 CCGCCGAGCTGAGGCGCAAGTTCTGTGAGAGAGCAGAGATCTTCAAGGAGTGAACAT 3553
 QY 1639 CCCAATGTGTGTGTCTATTTGTGTCTGACCCAGAGACCCCATCTTACATGTCTATG 1698
 DB 3554 CCCAATGTGTGTGTCTATTTGTGTCTGACCCAGAGACCCCATCTTACATGTCTATG 3613
 QY 1699 GAGCTGTGCAAGGAGGCGCACTTCTGTGACCTTCTCCGACCGGAGGGGCGCGCTGGCG 1758
 DB 3614 GAGCTGTGCAAGGAGGCGCACTTCTGTGACCTTCTCCGACCGGAGGGGCGCGCTGGAG 3673
 QY 1759 GTGAGACTCTGTGTGAGATGTGTGGGAGTGCAGTGTGTGCAATGAGATCTGTGAGAGC 1818
 DB 3674 ATGAAAGACTGATTAAGATGATGAGAAATGCGCGCGCGGCAATGAGATCTGTGAAAGC 3733
 QY 1819 AAGTGTGATTCACCGGAGACTGTGCTGTGAACTGTCTGTGTGAACAAGAAATGTC 1878
 DB 3734 AAGCACTGATTCACAGGAGACTGTGCTGTGCACTGTGTGTGACAGAAAGAACACC 3793
 QY 1879 CTGAATATCAAGTACTTGTGGAGTGTCCGAGAGGAGCCGATGGGGTCTATAGAGCTCA 1938
 DB 3794 CTGAATATCAAGTACTTGTGGAGTGTCCGAGAGGAGAGATGTGTCTATAGCTTCAAGC 3853
 QY 1939 GGGGAGCTGAGCAATGCCCGTGAAGTGAACGCACTGTGAGGCGCTTAACTACGCGCGC 1998
 DB 3854 GGGGAGCTGAGCAATGCCCGTGAAGTGAACGCACTGTGAGGCGCTTAACTACGCGCGC 3913
 QY 1999 TACTCTCCGAAAGCAGAGTGTGAGACTTGTGATTTGTCTGTGAGAGACTTCAAGCTTG 2058
 DB 3914 TACAGCTCCGAGAGGAGAGCTGTGAGACTTGTGAGATCTGTGTGTGAGAGACTTCAAGCTTG 3973
 QY 2059 GGGGCTTCCCGCTTATCTTCAAGTCTTCAAGCAATCAGCAACAGGAAATTTGTGAGAGAGGG 2118
 DB 3974 GGGGCTTCCCGCTTATCTTCAAGTCTTCAAGCAATCAGCAACAGGAAATTTGTGAGAGAGGG 4033
 QY 2119 GGGGCTTCCCGCTTATCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCA 2178
 DB 4034 GTGCGGCTGAGGCGCCGAGAGAGTGTGCGGAGAGATGTGTACGCTGTGAGAGAGCGCTGC 4093
 QY 2179 TGGGCTTATGAGCTGTGGAGAGCGGCCAGCTTACAGCAATCTTACAGAGAGCTTGCAGAGC 2238

Qy	2239	ATCCGAAAGCGGCATCG	2255	
Db	4154	ATCCGAAAGCGGCATCG	4170	
RESULT 10	AF033810			
LOCUS	AF033810	4788 bp	RNA	linear
DEFINITION	Fujinami sarcoma virus, complete genome.			
ACCESSION	AF033810			
VERSION	AF033810.1	GI:2801466		
KEYWORDS	Fujinami sarcoma virus			
SOURCE	Fujinami sarcoma virus			
ORGANISM	Vitruves; Retroid vitruves; Retroviridae; Alpharetrovirus.			
REFERENCE	1 (bases 1 to 4788)			
AUTHORS	Petropoulos, C.J.			
TITLE	Appendix 2: Retroviral taxonomy, protein structure, sequences, and genetic maps			
REFERENCE	(in) Coffin, J.M. (Ed.).			
AUTHORS	RETROVIRUSES: 757;			
TITLE	Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, NY, USA (1997)			
JOURNAL	2 (bases 1 to 4788)			
REFERENCE	Submitted (12-NOV-1997)			
AUTHORS	NTH, NLM, Rockville Pike, Bethesda, MD			
JOURNAL	20894, USA			
FEATURES	Location/Qualifiers			
SOURCE	1..4788			
REFERENCE	/organism="Fujinami sarcoma virus"			
AUTHORS	/mol_type="genomic RNA"			
TITLE	/db_xref="taxon:11885"			
JOURNAL	22..4442			
REFERENCE	/product="p140 polypotein"			
AUTHORS	22..102			
TITLE	380..3928			
JOURNAL	/codon_start=1			
REFERENCE	/product="p140 polypotein"			
AUTHORS	/protein_id="AAC82565.1"			
TITLE	/db_xref="GI:2801467"			
JOURNAL	/translation="MEAVIKVYISACKTYCGKTSPPSKKEIGAMLSDLOKEGLIMSLSD			
REFERENCE	LYSGSDMPITPAALTORAMVYKSGELTKWGLVYVGLAAAEBOVTSBOAKFMYGLGSLD			
AUTHORS	GRVSPGPECIEKATERERIDKGETVQDRIQALAAAEBOVTSBOAKFMYGLGSLD			
TITLE	CATSAAPPSPVYSGLYPSLAGVSEQGGDPTRGAAQPAEFGRTGLAAGPALTDBR			
JOURNAL	ARIEELASTGPEVNAFVYVITGEPATPTLEPKLIGDLAAGVAGRSFTAVAGVAV			
REFERENCE	ALMSPLPLPHDINPMRVILDPAPHALMMDMAASGOLHRPOBEHSTSLAAATWML			
AUTHORS	TOAESPHRLPHCSAASHODSHAMGPELMACBKHTELLRLDSERLLELMKKMML			
TITLE	SHAKSDREYAGMLHMFSSOLEKQEGHLEATDSOIGSMVNLASQTELLSOTLR			
JOURNAL	RAHELAAGPLAKLSTILIRDKOOLARKYSPBMOOLSOEVAWTOOEVEKTLAAOYRSIV			
REFERENCE	RDSYQARKYQOABKDEREAKKRYASLSKLYALNQVYLAQAAALHHNHVYQRA			
AUTHORS	LPTHEHSDYSLQDEVLVLKEILDEYCSITSVLEDVLAHQVAHVMETDPATERS			
TITLE	SFVTEHLYRSEVPPAYVFDESLLEAEMLLEGEQLNELTIESVOHSTSLIEBELAS			
JOURNAL	RKAVSSEQRWELQVELELGEELALSPERVHLKRGQGLAEAOQGLGVAQAKLC			
REFERENCE	MEYESKHCIRDLAARACVITEKNTLGLSDPGMSROBEDGVASTGGMKQIPYKMTA			
AUTHORS	PEALNYGWSSESDVMSFGILMEAFSLGAVPANYNLNQOTRAIEGVRLPEPEQCE			
TITLE	EDVRLNORCMEVYPHRRSPSCGAHODLIALRKHNR			
JOURNAL	380..835			
REFERENCE	/product="p14 MA"			
AUTHORS	899..1087			
TITLE	/product="p10"			
JOURNAL	1085..1309			
REFERENCE	/product="CA"			
AUTHORS	1307..3925			
TITLE	/product="FBS"			

ORIGIN	3'UTR	4196..4442
Query Match	45.1%;	Score 1018.2; DB 14; Length 4788;
Best Local Similarity	66.2%;	Pred. No. 2e-167;
Matches 1641;	Conservative	0; Mismatches 608; Indels 228; Gaps 3;
OY	1	ATGGGCTTCTCTTCTGAGCTGTGACAGCCCGCCAGCGGGGTCTCTGACCAATGACG 60
Db	1454	ATGGGCTTTTGGGCGCGAGCTGTGTGCGCCGAAAGGGGCACTGTGCTGTCTGGCTGGAG 15113
OY	61	GAGGCCGAGCTCGTCTACTGAGAGGAGATGAAAGTGAATGGCCAGCGGCTCAAGAGT 120
Db	1514	GACAGCGAGTTGGCCCTCTCTGAGCTGATGAAAGATGATGATCAGCGCTGTAAAGCC 15773
OY	121	GACAGGAGATATGAGAGATGTGCTTACACACATGTCC-----CTGCAGAGACAGT 168
Db	1574	GACCGGAGATACGGGGGGATGCTCAGCAACATGTCTCTGACGTGGAAGAACAGAGAGGC 1633
OY	169	GGGGGCCAGAGCGGGGCCATCAGCCCTTGACAGCCCCATCATGTCACTGCTGTGGCTGATGC 228
Db	1634	CTGGGACATCTCCGTGGCCACCAGCACAGCAGCCAGATCGGGAGATCGTGTGGGTTCTG 16939
OY	229	ACCAAGCCAACTGAGGGGCTGAGGCGCTGTGCGGGAGAGACGAGAGGATCTGAATCTCA 288
Db	1694	GCAAGCCAGATGAGACATGAGCCAGACGCTGTGGCGGACGCGGAGAGACTGGCAGCG 17553
OY	289	GGGCCCTTGAGCAAGCTGAGCTTCTCATCCGAGAAAGGCGAGCAGCTTCCGAAACTTAC 348
Db	1754	GGGCGCTGTGGCCAGCTGAGCATCTGATCCGTGACAAAGCAGCAGCTTCCGAAAGTCTTC 18133
OY	349	AGCAGCAGATGGGCGAGCTGCGAGCAGAGAGTCAACAAAGCCCAAGCCAGACGACATTGAG 408
Db	1814	AGTAGAGATGGGCGAGCAGCTCAGCCAGAGAGTACCGCTTGAGACCAACAGCAGAGAAATTGAG 18723
OY	409	AAGCTTGAAGACCGATACCGAGCTCTTGACACGGGACAGTGGCCCAAGCCAAAGCGAAGTAC 468
Db	1874	AAGCTGAAGGACAGTACCGGAGCTGTGTGTGTAACAGCAGCCAGGCGCAAGCGCAAGTAC 19333
OY	469	CAGAGAGCCAGCAAGAACAAGGACCTGTGACAAAGGCCAAAGGACAATGTATGTGCGAGCCTG 528
Db	1934	CAGAGAGCCAGCAAGACAGAGAGACGGGAAAGGCGAAGAAAGTACGTGTCCGAGCCTG 19939
OY	529	TGGAAGCTCTTGTGTCACCAACACCGGTATGTGTGGGCGTGGCGGCTGCGACGTACAC 588
Db	1994	TCGAAGCTCTATGTGCTTGCACCAACCAATGATGTGTGCTGTGGCAGGCGGCTGCTGCAC 20533
OY	589	CACCAAGCACCAACCAAGCTCTGTGCTCCCGGCTTGTGCGGTCACTGCAAGAACCTTGAC 648
Db	2054	CACCAACATCACTACACAGCGTGTCTGCCCACTGTGACAGATGCTCTTACAGCCTTGAG 21133
OY	649	GAGGAGATGGCTTGCATCTGGAAGAGATCTTGAGGAATPACTGAGAAATTAAGACCTG 708
Db	2114	CAGAGATGTGCTCTTGTGTAAGAGATTTCTTGGGAGTACTGTAGCATCAACAGTCTG 21773
OY	709	GTGCAAGATGAGTGTGTGCCATTACCGGAGATGTGCTGACGTGTGCCGACATCCAG 768
Db	2174	GTGCAAGGAGAGCGTGTGCCATCAACAGAAAGTTGCCAGCGCGCTGTGATGATGATGCAC 2233
OY	769	CTTGAAGCTGATGCCAAAGGCTTCTGTGACAGTATGGGTTCGCACTGACGTCCCAACC 828
Db	2234	CCCCCACTGATGATACACAGCTTGTGTTCAGTGTCCAGACAGTACGACTGTGAGTACACCG 22933
OY	829	TGTGTCAAGTTCATGATGATCTGCTTGGAAGAGGGTGAACCGCTGTGAGCTTGGGGAGCTC 888
Db	2294	GCAATGACCTTTGATGAGAGCTTGTCTGAGAGGCAAGAACTTGTGAGCCAGGGGAGCTG 23533
OY	889	CAGCTGAACGAGCTGATGTGTGAGAGCGTGTGACACACGCTGACCTCAGTACAGATGAG 948
Db	2354	CAGCTGAATGAGCTGACCATGTGAGAGGGGTGAACATCCCTGACATCATTTGAAGAGAG 24133
OY	949	CTGGCTGTGGCCACCGAGATGTGTCTTACAGCGGCGAGAGATGTGTTACGACCTGCAACAG 10088

Db 2414 CTGTTGCGCAGCAGAAAGGCGATGAGCAGCAAGCAGCGGTGTGGAGCTGACGATT 2473
 QY 1009 GAGCTCCGGAATGAAGAGAGAAACCCACCCCGGAGCGGTGAGCTGTGGCAAG 1068
 Db 2474 GAGCTGGGGGCGAGAGCTGCGATCAGCTTGGGGAGCGGGTGGACCTTGGCGAAG 2533
 QY 1069 AGGCAAGTGTGTCAGAAAGCACTGCAAGGGGCTGCAAGTACGCTGTGACAGCCAG 1128
 Db 2534 CGGCAAGGGGCTGGGGAGGCCAGCAGCAGCTGCAAGGCTTGGTCTGTGCTGAGCCAA 2593
 QY 1129 CTGCAAGCCCGCAGAGAGATTGCTGCAAGCCAAAGCTGAGACCTTGGGCCCGAGCC 1188
 Db 2594 CTGCAAGCCCGCAGAGAGATTGCTGCAAGCCAAAGCTGAGACCTTGGGGCTTGAAGCC 2653
 QY 1189 CGGCTGTGTGCTGCTGCGAGATGACCCGACCTCCAGCTGTGCTGTGAGCAGAGCGA 1248
 Db 2654 CTGTCGCCCTTGGCCCTTGCAGAGAGACCGGAGTCAAGCGCTTCACAGATCAGAGCCG 2713
 QY 1249 GAGGGGGGAGAGACACCCAGCTGTGAGATCTTAAAGCCACATCTCAGGAATCTTCCG 1308
 Db 2714 AGTGGGGTGA-----CTGCGCTGAAGACCATCAAGAACCACTTCGGGCATCTTCAGT 2767
 QY 1309 CCCAAGTTCTG----- 1320
 Db 2768 CCCAGGTTCTGCTGCCACCTCCGCTGCTCATCCGAGAGTGCAGAACCCGCTGTGC 2827
 QY 1321 ----- 1320
 Db 2828 CAGCAGGCTGTGATCCAGCGGGCCATCCGCGCTCGAGAGTGCAGAGCTGTGAATAC 2887
 QY 1321 ----- 1320
 Db 2888 AGCGGAGACTTCTGCGTGGCGGAGAGCCAGGGGAGAGAGATATGCTCAGCGTGTG 2947
 QY 1321 -----AACTGTATCCAGCTGGA 1338
 Db 2948 TGGGATGGGAGCGCCCGGCACTTCATCATCAGAGCTGTGACAACTGTACAGCTGGAA 3007
 QY 1339 GGGGAAAGGCTTCTAGCATTTCTTGTCTCATGCAACCACTACTAGAGCAACCCAGAGCC 1398
 Db 3008 GAGCATATGCTTCTCCACATCCGCTGTCTATGACCACTGCTGAGAGCCAGCGGCC 3067
 QY 1399 CTGACCAAGAGAGTGTGTGTCTGTCTGACAGGGCTGTGCCAAGACAAAGTGGTCTG 1458
 Db 3068 ATCACCCGGAAGAGCGGCATGTCTGTACAGAGCGGTGTCAAGGCAAGTGGTGTCTC 3127
 QY 1459 AACCTAGAGAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1518
 Db 3128 AACCATGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3187
 QY 1519 AGCGGAGCGCTGCGAGCGCAACAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1578
 Db 3188 AGCGGCGCTGT 3247
 QY 1579 CCACTGACCTCAAGGCGCAAGTTCTAAGAGAGCGAGATCTTGAAGAGTACAGCAC 1638
 Db 3248 CCGCGGAGCTGAAGGCGCAAGTTCTGTGAGAGCGAGATCTTCAAGAGTGAACAT 3307
 QY 1639 CCGCAATCTGT 1698
 Db 3308 CCGCAATCTGT 3367
 QY 1699 GAGCTGTGAGAGGGGGGAGCTTCTGACCTTCTGCGAGCGAGGGGGGGCGCGCTGCG 1758
 Db 3368 GAGCTGTGAGAGGGGGGAGCTTCTGAGCTTCTGCGAGCGAGGGGGGGCGCGCTGAG 3427
 QY 1759 GTGAGAGCTGT 1818
 Db 3428 ATGAGAGAGCTGAGTGAAGATGATGAGAAATGCGCGGCGGAGTGAAGTCTGTGAAGC 3487
 QY 1819 AAGTGTGTATCCACCGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1878
 Db 3488 AAGCACTGATCCACAGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3547

QY 1879 CTGAAGATCATGTAATCTTTGGATGTGTCCGAGAGAGAGCCGATGGGTCTATGACGCTCA 1938
 Db 3548 CTGAAGATCATGTAATCTTTGGATGTGTCCGAGAGAGAGATGTGTCTATGCTCTCAAG 3607
 QY 1939 GGGGGCTCTAGAACAACTCCCTGTGAATGATGACCGCAGCTGAGGCCCTTAACTACGCC 1998
 Db 3608 GGGGGCTCTAGAACAACTCCCTGTGAATGATGACCGCAGCTGAGGCCCTTAACTACGCC 3667
 QY 1999 TACTCTCCGAAAGGAGAGCTGTGAGCTTTGGATCTTGTCTGTGAGAGACTTACGCTTG 2058
 Db 3668 TACAGCTCCGAGAGGAGAGCTGTGAGCTTTGGATCTTGTCTGTGAGAGACTTACGCTTG 3727
 QY 2059 GGGGCTCTCCCTATCCCACTCAAGCAATCAAGCAGACACGAGAGTTTGTGAAGAGGG 2118
 Db 3728 GGGGCTCTCCCTATCCCACTCAAGCAATCAAGCAGACACGAGAGTTTGTGAAGAGGG 3787
 QY 2119 GGGGCTCTCCCTATCCCACTCAAGCAATCAAGCAGACACGAGAGTTTGTGAAGAGGG 2178
 Db 3788 GTGCGGCTGAGAGCCCGGAGAGAGTGTCCCGAGAGAGTGTATCGCTGTATGACAGCG 3847
 QY 2179 TGGGCTATGAGCTGTGGAGAGGGGCGGAGCTTCAAGACCATTAAGAGAGTGTGAGAG 2238
 Db 3848 TGGAGTATGAGCTGTGAGAGGGGCGGAGCTTCAAGACCATTAAGAGAGTGTGAGAG 3907
 QY 2239 ATCCGAAAGCGGAGATG 2255
 Db 3908 ATCCGAAAGCGGAGATG 3924
 RESULT 11
 M22820
 LOCUS 4901 bp DNA linear UNA 04-AUG-1993
 DEFINITION Figure 4. Nucleotide sequence of the proviral DNA of the variant
 clone of ST-FesV.
 ACCESSION M22820.1 GI:340749
 VERSION M22820.1
 KEYWORDS
 SOURCE unclassified
 ORGANISM unclassified
 REFERENCE 1 (bases 1 to 4901)
 AUTHORS Kemp,A.W., Bloemers,H.P. and Van de Ven,W.J.
 TITLE Structural analysis of a variant clone of Snyder-Theilen feline
 sarcoma virus
 JOURNAL Virus Res. 8 (4), 349-361 (1987)
 MEDLINE 88129048
 PUBMED 2829460
 FEATURES
 source Location/Qualifiers
 1..4901
 /organism="unclassified"
 /mol_type="unclassified DNA"
 /db_xref="taxon:12908"
 ORIGIN
 Query Match 37.3%; Score 841.4; DB 13; Length 4901;
 Best Local Similarity 78.3%; Pred. No. 1,2e-136;
 Matches 1122; Conservative 0; Mismatches 101; Indels 210; Gaps 1;
 QY 1033 ACCGACCCCGGAGCGGTGAGCTGTGGGAGAGGCAAGTGTCTGCAAGAGCACTG 1092
 Db 1874 ACCGACCCCGGAGCGGTGAGCTGTGGGAGAGGCAAGTGTCTGCAAGAGGCGCTG 1933
 QY 1093 CAGGGCTGAGGTATGAGCTGTGTGACCGAGGCCAACTGTGAGGCCAGAGAGTGTG 1152
 Db 1934 CAGGGCTGAGGTATGAGCTGTGTGACCGAGGCCAACTGTGAGGCCAGAGAGTGTG 1993
 QY 1153 CAGACCAAGCTGAGAGACCTGGGCGGCGAGCGCCCGGCTGTGTGTGTGTGTGTGT 1212
 Db 1994 CAGGCGAAGCTGAGAGAGCTGGGCGGCGAGCGCCCGGCTGTGTGTGTGTGTGTGT 2053
 QY 1213 GACCGCACTTCACGTGTCTTGTGAGCAGAGCGAGAGAGGGGGAGAGACCCACGCTG 1272

NRKSNPESODPVALTNLIESLVTHTOFTMDCCOOLLQALLTGERORVLEARKOV
PEEDRPQLPVNIDETPLTRBNMDFAPRAREHLRYROLLILGIRBAARPNTLA
QKOVYQCKEETPAAABADGTGFSSELSPOGHAEQWQWQELRLBEGKMAOR
VSDREYAGLLHMSLQDGGRGRTGYSPIQSMAETISQTEGLRLLQHEHNSG
PLSKGLILREKQRLKTYSEBMOQLOQLTHTNODLEKLSQYRALADQAARK
YQASDKDRDKAKLEGPBPVLLBLQDDRHSTSSSEORSEGRAPTELLKSHI
SGIFPKESLPRLDLPVEOKRHEHMLHGALPRAVAILTHSGDPLVSESGKO
EYVLSVMDGPRHRTIOSADNLYRBEQGPASIPLYNHLRSQOPLTKSGIYLR
AVFKRWLNHEDLVAGEQIGRNGFVSGRLRADNTLVAYSCHETLPIKAKFL
QAKILKOYSHPNIVLIGVCTQKOPITYIMELVQSGDELTELRTEGARLTKMLQ
VEDAAAGMEYLESKCIHVDLAARNCLVTEKVLKISDPSGDELVAASGGILQ
VEVKWTAPEALNYRYSSEVDWSEFGLIMETESLGAESPMLSEGNQTRFEVKGRL
PEPELCPDAVFLMEQCMAYEBGRPSFAITQELQSIKRRR"
mat_peptide
4.381
/product="protein p15"
mat_peptide
382.591
/product="protein p12"
mat_peptide
592.1044
/product="protein p30"
mat_peptide
1045.2871
/product="v-free transforming protein"
ORIGIN About 200bp upstream from a bgl11 site.
Query Match 35.9%; Score 810.8; DB 14; Length 2946;
Best Local Similarity 83.4%; Pred. No. 2.8e-131;
Matches 946; Conservative 0; Mismatches 182; Indels 6; Gaps 2;
QY 1129 CTGACAGCCAGCAGAGTTGCTGACAGACCAAGCTGAGACCTGGG---CCCCGGCAG 1185
DB 1738 CTGACAGCTGTAACAGAGGTGAGAGCCCTGACAGACAGCTGAGTGAACACGGGGCC 1797
QY 1186 CCCCCGCTGTCTGCTGACAGATGACCCGCACTCCAGTCTGCTCGAGACGAG 1245
DB 1798 CTCACAGGAGAGAGGTGCTGAGCTGTGACGCACTGAGGAACTTTCTGTGGCCGAG 1857
QY 1246 CGAGAGGGGGAGAGACCAACGCTGAGATCTTAAAGACCACTCAGAAATCTTC 1305
DB 1858 AGCCAGGGAGACAGAAATATGTCTGTGCTGTGTGGAGCGGCCAGCCCGCACCTTC 1917
QY 1306 CGCCCCAGTTCTC---GAACCTGTACCGACTGGAAGGAGGAGCTTCTAGCATTCCT 1362
DB 1918 ATCATCAAGTCCGCTGACAACTCTACCGAGGAGAGATGGCTTTCGAGCATCCCC 1977
QY 1363 TTGCTCATGACCACTTATGAGACCCAGACCCCTTACCAAGAAAGTGTGTGTTC 1422
DB 1978 TTGCTCTGACCACTGCTGTGCTCCAGACGCCCTTACCAAGAAAGCGGTATTTTC 2037
QY 1423 CTGACAGGGCTGTGCCAAGGACAGGTGGTGTGAACATGAGGACCTGGGTGGGT 1482
DB 2038 CTCACAGGGCTGTGCCAAGGACAGGTGGTGTGAACATGAGGACCTGGGTGGGT 2097
QY 1483 GAGCAGATTGACCGGGGAACTTTGGCGAAGTTCACGCGACGCTGCGAGCCGACAC 1542
DB 2098 GAGCAGATTGCGGGGGGGAACCTTTGGAGAGTTCACGCGACGCTGAGGGCCGACAC 2157
QY 1543 ACCCTGTGTGGCGTGAAGTCTTGTGAGAGAGGCTCCACCTGACCTCAAGGCCAAGTTT 1602
DB 2158 ACTCTGTGTGGCGTGAAGTCTTGTGCGGAGACCTCCACCTGACCTCAAGGCCAAGTTT 2217
QY 1603 CTACAGGAAGGAGAGTCTGGAAGAGTACAGCAACCCCAACATCGTGTGCTCATTTGT 1662
DB 2218 CTTCAGGAAGGAGAGTCTGGAAGAGTACAGCAACCCCAACATCGTGTGCTCATTTGT 2277
QY 1663 GTCTGACCCAGAGAGCCCATCTACATGTCATGAGACTTGTGAGGGGGGCGACTTC 1722
DB 2278 GTCTGACCCAGAGAGCCCATCTACATGTCATGAGACTTGTGAGGGGGGCGACTTC 2337
QY 1723 CTGACCTTCTCCGACCGAGGGGGGGCGCTGCGGGGTGAAGACTTGTCTGACAGATGTG 1782
DB 2338 CTGACCTTCTCCGACCGAGGGGGGGCGCTGCGGGGTGAAGACTTGTCTGACAGATGTG 2397
QY 1783 GGGAGTGCAGCTGTGAGTGAAGTCTGAGAGAGAGTGTGTCATCCAGCGGGAGCTCG 1842

DB 2398 GGGACGGCGCGCGCGGACGTAGAGTACCTGAGAGAAATGCTGATCCACGGGACCTA 2457
QY 1843 GCTGCTCGGAATGCTGCTGTGACAGAGAAATGTCCTGAAGTCAATGATCTTTGGGATG 1902
DB 2458 GCTGCTCGGAATGCTGCTGTGACAGAGAAATGTCCTGAAGTCAATGATCTTTGGGATG 2517
QY 1903 TCCCGAGAGGAAAGCGGATGGGCTTATGAGCTGACAGGGGGCTGACAGAGTCCCGTG 1962
DB 2518 TCCCGAGAGGAAAGCGGATGGGCTTATGAGCTGACAGGGGGCTGACAGAGTCCCGTG 2577
QY 1963 AAGTGACCGGACCTGAGCGCTTAATCTAGCGCCGCTACTCTCCGAAAGCAGCTGTG 2022
DB 2578 AAGTGACCGGACCGGAGCTTAACTAGCGCCGCTATCTCTGAGAGCAGCTGTG 2637
QY 2023 AGCTTTGGCATTTGCTGTGAGAGACCTTACGCTGGGGGCTTCCCTATTCACCTTC 2082
DB 2638 AGCTTTGGCATTTGCTGTGAGAGACCTTACGCTGGGGGCTTCCCTATTCACCTTC 2697
QY 2083 AGCATTCAGCAGACCGGAGTGTGAGAGGGGGGCGCTGCTGCCCGCCAGAGCTG 2142
DB 2698 AGCATTCAGCAGACCGGAGTGTGAGAGGGGGGCGCTGCTGCCCGCCAGAGCTG 2757
QY 2143 TGTCTGATCCGCTGCTTCAAGCTCAGTACAGAGCTGAGGCTGAGGCTGAGGCTGAGG 2202
DB 2758 TGTCTGATCCGCTGCTTCAAGCTCAGTACAGAGCTGAGGCTGAGGCTGAGGCTGAGG 2817
QY 2203 CCGAGCTTACGACCATCTTACAGAGAGCTGACAGAGCTTCCGAAAGCGGATCG 2256
DB 2818 CCGAGCTTACGACCATCTTACAGAGAGCTGACAGAGCTTCCGAAAGCGGATCG 2871
RESULT 14
BC073445 4788 bp mRNA linear VRT 03-AUG-2004
LOCUS
DEFINITION
Xenopus laevis MGC80946 protein, mRNA (cDNA clone MGC:80946
IMAGE:5516233), complete cds.
BC073445
ACCESSION
BC073445.1 GI:49116710
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 4788)
Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
Dev. Dyn. 225 (4), 384-391 (2002)
JOURNAL
PUBMED
REFERENCES
AUTHORS
2 (bases 1 to 4788)
12454917
Strausberg, R.L., Feingold, E.A., Grove, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schefer, C.F., Bat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Wax, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uadiri, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Miliady, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richard, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skaleja, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932
REFERENCE 3 (bases 1 to 4788)
AUTHORS Klein S. and Gerhard D.S.
TITLE Direct Submission
JOURNAL Submitted (14-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

REMARK
COMMENT

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabhu, Parveen Saeedi, UR Santos, Angeliq Schnerch, Ursula Skalska, Duane Smalins, Jeff Stott, Miranda Tsai, George Yang, Jacqui Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRK Plate: 155 Row: J Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES
source

Location/Qualifiers
1..4788
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:80946 IMAGE:551623"
/issue_type="Embryo, stage 31/32, Xenopus"
/clone_lib="NICHD_XGC_Emb4"
/lab_name="DH10B"
/note="Vector: pCMV-Sport6"
1..4788
/gene="MGC80946"
/db_xref="LocusID:444292"
399..2867
/gene="MGC80946"
/codon_start=1
/product="MGC80946 protein"
/protein_id="AAH73445.1"
/db_xref="GI:49116711"
/db_xref="LocusID:444292"
/translation="MGFAELQCPNGHATLLRIQDTLAVMSMKCTORAKDKKEY
SNNLHQISVQEKLDQSLTPEGLSEYSSQSESWATLVSTENISQILKRSSEVNA
LSLTLIREKQDKKSYSEOWQLNODPMKTOODIEKLCYRQVETQSKKX
OEACRDKDKAKKVKNTWKHLAHNVLAVRPAQDHEHHPVSLPTLHESQS
HIOEWQIILKNIIOEYSITSLVQSEVNVHREIASAIVSEQEDNPLQMSKVS
EUPRYTPPTSLIEETDLEBEQLNDLVYESVQHTLVSEBLSAIVTAISKQD
KQHSIDIAIEKTDRCBQNLAKRHARDALQYHMTLCNKKIQGRREMLKRMN
TIGFDPSPALQIDNDSITSVRECGKKTLENIKNHISGIFRPFPSPLPLV
VQPLSQQNMWGAIPRSEVOGLVNRGFLVRESQKQBYLVSLVMDQPRHFI
VNLVYLEEGESTIPLLINHPVKTQAAVTKSGVINKAIYDKRMVEEYILGER
IGKNGFVSESGRLADNTPVAVKSCDRLPDLDKRPMEARILKQYHNPVTKIG
VCGQRKPIYVMEIYVGGPQPTLQNEGRILYKELIRVSEAAAAMEYLSHCIR
DLAARCLVTERKNAKISDPGMSREDEYISTGSKQKPIPKMTAPALANTGRYS
SDVMSFGILMEAFSLGSVPYAAMTNQCTREALDQGVRLVDPNCPDEVSLMLRCWE
YPKKRPNSIVQVAVTIRKRYR"

ORIGIN

Query Match 29.1%; Score 656.4; DB 5; Length 4788;
Best Local Similarity 57.6%; Pred. No. 2e-104;
Matches 1421; Conservative 0; Mismatches 831; Indels 216; Gaps 5;

QY 1 ATGGGCTTCTCTTCTGAGCTGTGACGCCGCCAGGGGCTCTGACGAATGACG 60
DB 399 ATGGGGTTGGCGGAAGATTACAGTGGCCATGAGCTTATGTTACGTTACAA 458
QY 61 GAGGCCGAGCTTCTGTCTACTGAGGGCATGAGAAATGATGAGCCAGGGGCTCAAGAT 120
DB 459 GACACTGGAATTGAGGTGATGAGTGCATGAAATAATGTTTATATACAGGCTGCAAGGGG 518
QY 121 GACAGGAGTATGACGAATGCTTCCACCATATGTCCTGACAGG---ACAGTGGGGGCA 176
DB 519 GACAAAGATATTCGAATATGCTTACACGAATTTCTGTGACAGTGCAGAAAAGCTGATCAG 578
QY 177 GAGCCGGGCTATCAGCCCTGAC-----AGCCCATGTCAGTCTTGGGCTGAGATC 228
DB 579 TGGCTGACCCCTGAGTATATCTGAATCTCAACAGCTCAGTGAAGTCTTGGGACAGTTG 638
QY 229 ACCAGCCAACTGAGGGCTGAGCCGCTTGTCTGCGGACACACGAGAGATCTGAATCTCA 288
DB 639 GTGTCTCAGACAGAGAACTTAAGCCAGATTTTACGGAACATTCAGAGAGATGAATGCA 698
QY 289 GGGCCCTGAGGAGGCTGAGCTCATCCGGGAAACGAGAGCTTGCAGAGCTTAC 348
DB 699 GGAACCTGAGCAAAATTAACATTTCTTATCCGGAAAGCAGCAGCTTAAAAAGTCTTAT 758
QY 349 AGCGAGCAGTGGCAGCAGCTGACAGAGAGCTCACCAGAACCCACAGCAGCATTTAG 408
DB 759 AGTGAAGCAATGGCAATCTTCAACCAAGATTCATGAGAACACACAGAGATATCGAA 818
QY 409 AGCTGAGAGCCAGTACCGAGACTCTGCAACGGACAGTGGCCCAAGCCAGCAATGAC 468
DB 819 AAGCTGAGGTGTGAGTATGCGCAAGCAAGTAAAGAAATCTTCCMACTCAAGAAATAATAC 878
QY 469 CAGAGAGCCAGAAAGACAGGACCGTGAAGGCAAGCAAGATATGCGCAGCTTG 528
DB 879 CAAAGAGCATGAGAGATTAAGACCGCATTAAGCAAGAGAGATGTTGAAGATAC 938
QY 529 TGAAGCTCTTGTCTCACCACAAACCGATATGTGTGGGCTGCGGCTGCGCAGTTAC 588
DB 939 TGGAAATCATGACATGATTAACCATATGTCCTGTGTTAAGAACTGCTCAGTTAC 998
QY 589 CACACAGCACACACACAGCTCCTGCTGCGCGGCTGCTGCGGCTCATGAGAGACCTGAC 648
DB 999 CATGAGCACACATTTCTGTATAGTCTGCCAACATGCAAGATCCATGAGACATTCAC 1058
QY 649 GAGAGATGCTTGTGATCTGAAAGAGATCTGCAAGAAATACCTGAGATTAAGCACTTG 708
DB 1059 CAGAGATGCTCAAGATTTCTCAAGAAATATTAAGCAAGATATCTGAAATCACAATCTTA 1118
QY 709 GTGCAGATGAGGTGTGCAATTCACCGGAGATGCTGCAAGTCTGCGCGCATTCAG 768
DB 1119 GTACAAAGAGAGTATGTTCCCTGCAATAGAAATGGCAGTGTGTTAAGAAATCAGC 1178
QY 769 CCTGAGGCTGAGTACCAAGGCTTCTGCGACAGTATGAGTCTGCACTGACCTCCACCC 828
DB 1179 CCTGAGCAGAAATATGCAATTTCTCCAGATGAATAAGGTGCTGATTAATGCCACCC 1238
QY 829 TGTGTACGTTGAGATGACATGCTTGTGAGAGGGTGAACCGCTGAGAGCTGAGAGCTC 888
DB 1239 ATTGTGACATTTGATTCAGATTTGCTGAGAGAGACAGATGACTGGA---AGAAAGACTTA 1295
QY 889 CAGCTGAAGAGCTGACTGTGAGAGCGTGCAGACAGCCTGACCTCACTGATGACAGTGA 948
DB 1296 CAGCTGAATGACCTGACATGTAAGATGTGCAATACCTTGAAGCTGTTGAAGAGAA 1355
QY 949 CTGGCTGTGCGCACCGAGATGAGTGTTCAGAGCGGACAGAGATGTTACGACGCTGCAACAG 1008
DB 1356 CTGTCTTCACTACAGCGCGGCATGGAACCAACAGAGATATAGTGAACAACTACCTCC 1415
QY 1009 GAGCTCCGAGATGAGAGAGAACCCACCCCGGAGAGGAGTGCAGCTGCTGGGCAAG 1068
DB 1416 GATATTGCAATGAGAAAGAACACACAGATTTGACAGAAACAGCTTATCTGTTGGCCAA 1475
QY 1069 AGGCAAGTCTGCAAGAACATGCACTGACAGGGGCTGACAGTGAAGCTGTGACAGCCAGGCCAG 1128

Db	1476	AGACATGCTACGGGATGCCCTTCTGCAATATCAATGACTCTCTGCAACAAGAACAG	1535
Oy	1129	CTGCAGGCGCCAGCAGAGATGTGCTGCAGACCAAGCTGAGCACCTGGGCCCGGAGGCC	1188
Db	1536	ATCCAGGCGCCAGAGAGAAATGCTAAAGAGAAAGATGAACACTATTGGATTGGGGACCTT	1595
Oy	1189	CCGCTGTGCTGCTCTCTGACGAGATGACCGCACTTCCAGTGTCTCTGACAGAGACGA	1248
Db	1596	CCTCCTGCGCTTCAACTTCAGAGATGATTAACAGTCAATTACTTTCAGTGGACGGGAATGT	1655
Oy	1249	GAGGGGGAGAGACACCAACGCTGAGA-----	1276
Db	1656	GGAAGAAAACATTAGAAAATATAAAAATCAATTTCTGGAATATTTCCGCTAAATTT	1715
Oy	1277	-----TCTTTAAGAGCACATCTCAGAAATCTTCGCGCCAGATGCTC	1319
Db	1716	TCTTTGCTCCTCTTTTATCTTTGTGTCACAGATGTCCAGAACTCTCAGCCMACAGAC	1775
Oy	1320	-----	1319
Db	1776	TGTAACATGATGCCATACCTCGCTGAAAGTCAAGACTGTTGTGAAATCGAGGGAT	1835
Oy	1320	-----	1319
Db	1836	TTCTCTTGTCCGGAGAGCCAGGGCAGAGAAATATGCTGTCTGTACTATGGATGGG	1895
Oy	1320	-----GAACTGTACCGACTGGAAGGGAGGC	1347
Db	1896	CAACACGACATTTTATATCATTCAGACGTAGATATCTTTACAGTTTGAAGAGAGAGGC	1955
Oy	1348	TTTCTTACATTCCTTTTGTCTCATATGACCACTACTGACACCCAGCCCTCCACCAAG	1407
Db	1956	TTTTCAAGATCCCAATGCTGTCAATCACTTTGTCAAAACCCAGCAAGCATGTGACAAA	2015
Oy	1408	AAGATGTGTGTTCCTGSCACAGGCGTGTGCCAAAGACAAAGTGTGTCTGAACATAG	1467
Db	2016	AAGATGTGTATATCATTAATTAAGCTATTGTTAAGGCAAAATGCTATGAGACTGAA	2075
Oy	1468	GACCTGGTGTGGGTGAGCAGATTGGAAGGGGGAACTTTGCGAAGTTCACGGGAGCG	1527
Db	2076	GAACTCATCTGGGAGAGAAATTTGAAAGGGAATTTTGTGAATATTCACGCGAGCG	2135
Oy	1528	CTGCGAGCCGACAAACACCTGTGTGCGGTGAAGTCTTGTGACAGAGACGCTCCACTGAC	1587
Db	2136	CTTGGGCGTGAATACACACACAGTGGCAATTAAGAGCTGTCCGGACACACCTCCGCAAC	2195
Oy	1588	CTCAAGCCCAAGTTTCTACAGAAACGAGATCTGAAGCAATACGCCAACCCCAATC	1647
Db	2196	TTAAAGGCTAAGTTTCTAATGGAAGCCGAATACTGAACAAATATATGCAACCTTAACTT	2255
Oy	1648	GTCGGCTCATTTGGTGTGACACCCAGAAAGCAGCCATCTACATGTGATGAGACTTGTG	1707
Db	2256	GTCAAACTTATTTGAGTCTGCACACAAATACTCAATTTTACATTTATGAGTATGTG	2315
Oy	1708	CAGGGGCGCACTTCTCTGACCTTCTCCGCAACGAGAGGGGCGCGCTCGGATGAAGACT	1767
Db	2316	CAAGGTGAGACTTTCACACGTTTCTACAGATGAAAGTCCAGGCTGAAGTGAAGAG	2375
Oy	1768	CTGCTGCAGATGTGGGAGATGACAGCTGTGCAATGAGTACCTGAGAGCAAGTCTGC	1827
Db	2376	CTTATTAAGGTGTCAAAAATGCTGCGGCTGGAAATGAGATATTGGAGATTAACATTTGT	2435
Oy	1828	ATCCACCGGGACCTGGCTGTCCGAACTGCGCTGGTGAAGAGAAATGTCTGAAGATC	1887
Db	2436	ATTCACAGGGACCTTGACAGCCCGTAACTGTCTGATCTGAGAAAAATATGCTTAAAGATA	2495
Oy	1888	AGTGACTTTGGGATATCCCGAGAGAAACCGATGGGATCTATGCAAGCTCCAGGGGCGCTC	1947
Db	2496	AGTGACTTTGGCATATCCAGAGAAATAAAGATGAGATGATATCACTACATGGGGGATG	2555
Oy	1948	AGACAAATCCCCGTGAAGTGAACGCAACTGAGGCCCTTAACTACGCGCGCTACTCTCC	2007

Db	2556	AAGCAGATTCTTAAATATGAGCGGGCCCGAGAAAGCCTCAATTACGGTCGTACAGTTCA	2615
Qy	2008	GAAGCAGACGTGTGAGACCTTTGGCATCTTGTCTGTGGAGACCTTCAGCCTGGGGGCTCC	2067
Db	2616	GAGAGTATGTTTGGAGTTTGGTATCTTACTCTGGAGAGCTTCAGCTGGGTTCTGTT	2675
Qy	2068	CCCTATCCCACTCAGCAATCAGCAGACACGGGAGTTTGTGGAGAAAGGGGGCCGTCTG	2127
Db	2676	CCTTATCAGCTATGAGCAACCGAGCAAAACAAGGAGGCCATAGAAACAGGGGAGTTTCCTG	2735
Qy	2128	CCCTGGCCAGACGTGTGTCCTGATGCGCGGTGTCAAGGCTCAATGAGCAGTGTGGGCTCAT	2187
Db	2736	CTTGTCCCAAGTAATCTCCCTGATGAAGTCTATAGCTTGAATGCTGAGATGCTGGAAATAT	2795
Qy	2188	GAGCCTGGGCGACGGCCAGCCTTCAGCAGCATTCACAGAGAGCTGCAGAGCATCCGAAG	2247
Db	2796	GATCCAAAGAACGGCGCTTAATTTCAGCATTTGTGATCAGGTCCTTTCACCATCCGAAAG	2855
Qy	2248	CGGCATCG 2255	
Db	2856	AGATACCG 2863	
RESULT 15			
A06935			
LOCUS	A06935	856 bp	RNA linear PAT 04-OCT-1993
DEFINITION	fes/fps-related mRNA.		
ACCESSION	A06935		
VERSION	A06935.1	GI:490473	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1 (bases 1 to 856)		
AUTHORS	Van de Ven, W.J.M., Roebroek A.J.M. and Schalken, J.A.		
TITLE	Recombinant DNA and cDNA, mRNA, protein, antibodies, and a method		
JOURNAL	of detecting tumor cells		
FEATURES	Patent: BP 0246709-A 2 25-NOV-1987;		
source	Stichting Katholieke Universiteit		
	Location/Qualifiers		
	1..856		
	/organism="synthetic construct"		
	/mol_type="unassigned RNA"		
	/db_xref="taxon:32630"		
	3..638		
	/note="Protein sequence is in conflict with the conceptual		
	translation"		
	/codon_start=1		
	/transl_table=1		
	/product="fes/fps-related protein"		
	/protein_id="CAA00604.1"		
	/db_xref="GI:4526313"		
	/translation="KQYSHPNIVRLIGVCKOKQIYIVMELNOSGDFLPLRTGARGL		
	RVKTLVLWVGDAAGMELYSKCCITHDLARNCLVTERKAVLKISDPGMSRENDQIY		
	AASGSLNQVPMKTAPALNYGRYSSSDVMSFGLIMETFSGLASYFPMISNOOTRE		
	FVEKGRLLCPBELCPDAVFLMEQCMAYBPGORPSFTIYQELQSIKRRH"		
ORIGIN			
Query Match	28.1%	Score 635;	DB 6; Length 856;
Best Local Similarity	100.0%;	Prod. No. 1,4e-100;	Indels 0; Gaps 0;
Matches 635;	Conservative 0;	Mismatches 0;	
Qy	1682	CCATCTACATGTCATGAGAGCTTGTGACGAGGGGCGACCTTCTGACCTTCTCCGCAAG	1741
Db	61	CCATCTACATGTCATGAGAGCTTGTGACGAGGGGCGACCTTCTGACCTTCTCCGCAAG	120
Qy	1742	AGGGGGCCCGCTTCGGGTGAAGCTTCTGTCGACGAGTGTGGGGGAGTGCAGCTGTGGCA	1801
Db	121	AGGGGGCCCGCTTCGGGTGAAGCTTCTGTCGACGAGTGTGGGGGAGTGCAGCTGTGGCA	180

```
OY 1802 TGAAGTACCTGAGAGCAAGTCTGATCCACCGGACCTGGCTGCTCGAACTGCTGG 1861
    |||||
DB 181 TGAAGTACCTGAGAGCAAGTCTGATCCACCGGACCTGGCTGCTCGAACTGCTGG 240
    |||||
OY 1862 TGACAGAGAAAGATGTCTGAAGATCAGTGAATGTTGGATGTCCGAGAGAAAGCCGATG 1921
    |||||
DB 241 TGACAGAGAAAGATGTCTGAAGATCAGTGAATGTTGGATGTCCGAGAGAAAGCCGATG 300
    |||||
OY 1922 GGGTCTATGCAAGCTTCAGGGGGCTTCAGCAAGTCCCGGTGAAGTGACCGGACCTGAGG 1981
    |||||
DB 301 GGGTCTATGCAAGCTTCAGGGGGCTTCAGCAAGTCCCGGTGAAGTGACCGGACCTGAGG 360
    |||||
OY 1982 CCTTAACAGGAGCGCTACTCCTCCGAAAGCAGAGTGTGAGACTTTGGCATCTTGCTCT 2041
    |||||
DB 361 CCTTAACAGGAGCGCTACTCCTCCGAAAGCAGAGTGTGAGACTTTGGCATCTTGCTCT 420
    |||||
OY 2042 GGGAGACCTTCAGCCTGGGGGGCTCCCGCTATCCCAACTCAGCAATCAGCAGACAGGG 2101
    |||||
DB 421 GGGAGACCTTCAGCCTGGGGGGCTCCCGCTATCCCAACTCAGCAATCAGCAGACAGGG 480
    |||||
OY 2102 AGTTTGTGAGAAAGGGGGGGCCGTGCTGCCCAAGAGCTGTGTCTCTGAATGCCGTGTA 2161
    |||||
DB 481 AGTTTGTGAGAAAGGGGGGGCCGTGCTGCCCAAGAGCTGTGTCTCTGAATGCCGTGTA 540
    |||||
OY 2162 GGCTCATGAGCAGTGTGGGCTATGAGCTGTGGCAGCGGCCAGCTTCAGCACCATCT 2221
    |||||
DB 541 GGCTCATGAGCAGTGTGGGCTATGAGCTGTGGCAGCGGCCAGCTTCAGCACCATCT 600
    |||||
OY 2222 ACCAGAGCTGCAGAGCATCCGAAAGGGGCATGG 2256
    |||||
DB 601 ACCAGAGCTGCAGAGCATCCGAAAGGGGCATGG 635
    |||||
```

Search completed: March 19, 2005, 10:14:01
Job time : 9411.68 secs